Writing and Language Learning

Advancing research agendas

Edited by
Rosa M. Manchón

Writing and Language Learning

Language Learning & Language Teaching (LL&LT)

The LL&LT monograph series publishes monographs, edited volumes and text books on applied and methodological issues in the field of language pedagogy. The focus of the series is on subjects such as classroom discourse and interaction; language diversity in educational settings; bilingual education; language testing and language assessment; teaching methods and teaching performance; learning trajectories in second language acquisition; and written language learning in educational settings.

For an overview of all books published in this series, please see *benjamins.com/catalog/lllt*

Editors

Nina Spada
Ontario Institute for Studies in Education
University of Toronto

Nelleke Van Deusen-Scholl
Center for Language Study
Yale University

Volume 56

Writing and Language Learning. Advancing research agendas
Edited by Rosa M. Manchón

Writing and Language Learning

Advancing research agendas

Edited by

Rosa M. Manchón
University of Murcia

John Benjamins Publishing Company
Amsterdam / Philadelphia

Table of contents

Part II. Advances in empirical research

Part III. Advances in future research agendas

Coda

Introduction

Writing and language learning
Looking back and moving forward

Rosa M. Manchón

University of Murcia

This introductory chapter serves two main purposes. One is to contextualize the book within the larger professional discussion, which entails a look back approach in order to provide a synthetic review of key milestones and developments in the study of writing as a site for language learning. The second aim is to introduce readers to the aims, structure, and contents of the book. To this end, I (a) outline the key elements of the various contributions to the book and the way in which they are framed in a common research agenda; (b) describe the interconnection among contributions and the internal coherence of the volume; and (c) advance the way in which the book attempts to move theory and research forward.

The inquiry into writing as a site for language learning is a newcomer to second language acquisition (SLA) studies, which, less than ten years ago, was characterized as "a well defined space for a *future* research domain at the intersection between L2 writing and SLA" (Manchón, 2011a, p. 62. Emphasis added). Since then, theoretical and empirical work has grown exponentially and the "future" mentioned in the quote has distinctively evolved into the "present". As a result, the past 10 years have seen substantial progress in answering questions about how and why the acts of writing and feedback appropriation can lead to learning, not only writing but also language. In this way, the SLA field has come a long way from the time when, as Linda Harklau wrote at the turn of the century, the role of literacy in SLA was overlooked in Applied Linguistics research and, as a consequence, reading and writing were "peripheral concerns in studies of second language acquisition in classroom settings" (Harklau, 2002, p. 335). The "modality-sensitive" perspective that Harklau cogently advocated as being theoretically- and pedagogically-relevant for instructed SLA studies has permeated research agendas to such an extent that at this point in the development of this area of inquiry it is even possible to adopt a retrospective look and assess critically

the advancements made, as attested by recent narrative reviews and meta-analyses of the most outstanding research in this body of work (Bitchener & Storch, 2016; Manchón & Vasylets, 2019, for representative examples).

The current volume aspires to add to these previous initiatives from a dual perspective: It seeks to reflect current progress in the domain as well as to foster future developments in theory and research. To achieve this dual aim, the book contributes a collection of theoretical and methodological reflections about present and future research agendas, as well as a body of new empirical evidence intended to advance current understandings of the theoretically predicted language learning affordances of L2 writing. The theoretical postulations contained in Part I identify and expand in novel ways the diverse lenses through which the varied, multi-faceted dimensions of the connection between writing and language learning can be explored. The methodological reflections put forward in Part III signal to theoretically-grounded and pedagogically-relevant paths along which future empirical work can grow. The empirical studies reported in Part II, framed in a common research agenda, exemplify the diverse theoretical paradigms that can be applied to the study of writing as a site for language learning. They zoom into the many facets of the phenomenon and, collectively, illuminate the myriad of individual, educational, and task-related variables that (may) mediate short-term and, importantly, long-term language learning outcomes. These studies examine diverse forms of writing, performed in varied environments (including pen-and-paper and digital writing), conditions (writing individually and/or collaboratively), and instructional settings (academic settings – including secondary school and college level institutions – as well as out-of-school contexts).

This introductory chapter serves two main purposes. One is to provide readers with a synthetic review of key milestones and developments in this strand in order to contextualize the book within the larger professional discussion. Against this backdrop, I shall provide a preview of the structure and contents of the book. Readers should turn to the Coda chapter (Chapter 17) for an assessment of the manner in which the contributions to the book collectively enhance research insights and point to future directions capable of deepening current understandings of L2 writing as a site for language learning.

**Looking back: A synthetic review of key developments in theory
and research on writing and language learning**

Cumming's pioneering work: A writing perspective

All accounts of L2 writing as a site for language learning acknowledge Alister
Cumming's seminal work on writing processes as the original impetus for the theo-
retical and empirical interest in the connection between writing and language learn-
ing. Thus, in the early nineties, Cumming (1990) reanalyzed and reinterpreted part
of the think-aloud data he had collected for his PhD thesis (published in Cumming,
1989) from 23 adult Francophone learners of English while they composed L2 texts
of different complexity. On the basis of this empirical evidence and his interpreta-
tion of it, Cumming (1990) offered the first formulation of L2 writing as a site for
language development, which he articulated as follows:

> Composing might function broadly as a psycholinguistic output condition wherein
> learners analyze and consolidate second language knowledge that they have pre-
> viously (but not fully) acquired […] Composition writing elicits an attention to
> form-meaning relations that may prompt learners to refine their linguistic expres-
> sion – and hence their control over their linguistic knowledge – so that it is more
> accurately representative of their thoughts and of standard usage.
>
> (Cumming, 1990, p. 483)

Importantly, Cumming associated the learning potential of writing with com-
municative events in which the writer is fully involved in an intense linguistic
meaning-making activity, a perspective very much in line with Byrnes's theoretical
postulations in Chapter 4 (this volume), especially her enforced argument for the
key role that the meaning-making activity that is criterial to writing possesses in
bringing about language learning gains. In these writing conditions, Cumming
originally observed, L2 writers engage in thinking episodes characterized by an
intense meaning-making activity whereby they pay attention to both "the substan-
tive content of a text and its linguistic constituents while composing it" (p. 504).
Crucially, he contended that these "thinking processes may be more effectively
fostered when language learners are prompted to exert intentional control over
their own written expression" (p. 504), therefore pointing to intentional linguistic
processing while composing as the necessary condition for writing to result in
language learning gains.

This proposal, closely linked to the heightened attention to language in the writ-
ing condition emphasized in recent postulations of writing as language learning (e.g.
Manchón & Williams, 2016; Williams, 2012), is revisited in his own contribution to

the present volume (Chapter 2), where it is reiterated that writing leads to language learning "through processes of composing", observing that "conspicuous opportunities for learning the L2 appear when writers evaluate forms of the L2 in relation to their intended meanings, search earnestly to find the best words to express ideas, and switch purposefully between languages to make principled decisions" (pp. 32). Readers will encounter further elaboration of the anticipated language learning benefits of this intense linguistic processing activity in Chapter 4 (Byrnes), in some of the empirical studies reported in Part II (see especially Chapter 10 by López-Serrano, Roca de Larios, & Manchón), and in the suggestions for future inquiry into writing processes from the perspective of their language learning potential put forward in Chapter 14 (Manchón & Leow).

Despite the substantial amount of SLA-oriented L2 writing research over the last years, research agendas have not made a priority of the identification and description of those composing processes in individual writing conditions that Cumming anticipated to "have potential for learning of the language" (Cumming, 1990, p. 484). Therefore, many issues relating to the language learning affordances of the processing dimension of writing remain unresolved. This important gap in research justifies arguments in favor of an exploratory, process-oriented L2 writing research agenda (see Manchón & Leow, Chapter 14, this volume) as a much needed line of inquiry whose insights can eventually lead to the kind of experimentation that Cumming envisaged in 1990, when he stated that his own study simply aimed to "establish descriptive data preliminary to conceiving research that would be capable of doing so through experimentation or controlled longitudinal studies" (p. 488).

Cumming also stressed the joint forces of the metalinguistic and ideational elements of composing, claiming that writers focus on language "while concurrently thinking about their ideas and semantic meanings" (Cumming, 1990, p. 500), a prediction that has not received sufficient empirical attention thus far (but see Byrnes, 2013, 2014) and one that readers will see theoretically and empirically addressed in several chapters in the current volume (see especially Chapters 4, 10, 12, and 13).

Linda Harklau's contribution: An instructed SLA perspective

The next building block in pushing SLA-oriented L2 writing research agendas forward is unmistakably represented by Linda Harklau's study cited in the opening paragraphs and published twelve years after Cumming's (1990) pioneering work: "The role of writing in classroom second language acquisition" (Harklau, 2002). Similar to how Cumming's postulations developed, Harklau's positioning on the role of writing in instructed L2 learning also resulted from her own empirical research. Worthy of note is that in this case we are not concerned with the (controlled)

study of writing processes, or with adult, college-level writers (as was the case in Cumming's, 1990 study), but rather with classroom-oriented research conducted in secondary schools in US. This is a relevant point because, as will be noted at different points in the chapter, classroom-based research has only recently made its way into SLA-oriented L2 writing research agendas.

Harklau's classroom observations distinctively evidenced the dissociation between, on the one hand, the assumed prominence of oral language in language classrooms and its central role in any L2 learning that may derive and, on the other hand, the prominence and relevant role of literacy practices in the secondary school language classrooms she observed. In her own words: "my research in American public schools has convinced me that reading and writing are of relevance to virtually all classroom-based research" (Harklau, 2002, p. 330). She therefore cogently argued for a more central position of the study of the language learning potential of writing in both L2 writing and instructed SLA research agendas. Interestingly, Harklau's position paper was published in the L2 writing flagship journal, the *Journal of Second Language Writing* (*JSLW*), which at the time was primarily oriented towards composition writing and, as a result, links with SLA studies were minimal. This situation has changed remarkably over the years and the instrumental role played by *JSLW* in driving this strand of research forward ought to be duly acknowledged (see below).

Harklau's contribution represented a fundamental building block from a dual perspective. From a writing research angle, her often cited admonition that "while it is important for classroom-based studies to investigate how students learn to write in a second language, it is equally important to learn how students learn a second language through writing" (p. 329) in effect constituted the first call for a much needed and welcome widening of the scope of writing studies to include the two dimensions of "learning-to-write" and "writing-to-learn" (Manchón, 2011a) as equally legitimate and relevant research foci. From an SLA research perspective, her claim that "writing should play a more prominent role in classroom-based studies of second language acquisition" (p. 329) equally constituted a far-reaching postulation that may not have been sufficiently acknowledged in the field: It provided a strong foundation for redressing the oral-bias of SLA studies, a well attested and widely acknowledged fact at present (see, for instance, Byrnes & Manchón, 2014a,b), although not at the time when Harklau's paper was first published. In effect, she finished her article categorically stating that the L2 writing community had "two legitimate and vital roles to play" (p. 345), one of which she envisaged as entailing no more and no less than "to interrogate research and theories of second language acquisition that do not adequately account for the role of literacy in classroom learning" (p. 345). Accordingly, Harklau's work ought to be seen as a key stepping

stone in elevating the status of the study of writing to a central position in instructed second language acquisition (ISLA) research agendas.

Yet, despite notable attempts in this direction (see chapters in Part III), the bulk of research on writing as language learning has been framed primarily in SLA rather than ISLA postulates, with the result that ISLA-oriented classroom-based studies do not abound. The current volume decidedly seeks to redress this situation: It includes educationally-oriented, ISLA-informed theoretical and research methodology reflections (Byrnes, Chapter 4; Leow, Chapter 5; Manchón & Leow, Chapter 14), as well as a set of longitudinal, classroom-based empirical studies (see especially Chapters 7, 12 and 13) that shed new light on Harklau's predictions regarding the "the instrumental role that writing can play in the acquisition of a second language in educational settings" (p. 345). It is also worth noting that diverse educational settings are in focus in the classroom-based empirical studies included in the current volume: an out-of-school context (Chapter 7), a university English language teaching program (Chapter 12), and a secondary-school CLIL program (Chapter 13).

Manchón and Roca de Larios's contribution: A problem-solving, SLA-oriented, L2-writing perspective

Chronologically, a third building block in advancing disciplinary discussions in the domain was represented by a contribution from two L2 writing scholars with an SLA background and working in a foreign language context. I am referring to Manchón and Roca de Larios's (2007) position paper "Writing-to-learn in instructed language learning contexts", regarded as "the first formal appearance of WLL [writing to learn language] as a specific dimension for L2 writing" (Ortega, 2012, p. 240). As with Cumming's 1990 and Harklau's 2002 studies referred to earlier, our reflections on the connection between writing and language learning originated in our own empirical research, namely, our sustained program of research on writing processes with L2 writers at different proficiency levels while writing in their L1-Spanish- and L2 -English (see overview in Manchón, Roca de Larios, & Murphy, 2009).

It is of relevance to note that, as was the case with Cumming's work, our research on writing processes was not originally linked to an interest in writing as a site for language learning. Rather, our global aim was to contribute to L2 writing theorizing and to do so with empirical evidence on the cognitive dimension of composing provided by writers learning and writing in a foreign language context, a setting less visible in L2 writing studies at the time (but see Manchón, 2009, for a collection of sustained research programs on writing in foreign language contexts

in diverse geographical locales). Nevertheless, it is important to note that although we were not primarily concerned with the language learning affordances of writing, we did make a priority of the formulation process in our research program. This was because our central concern was to inspect our participants' attempt to transform ideas into language, and to zoom into the anticipated idiosyncratic nature of the problem-solving behavior that such conversion of ideas into language could possess while writing in an additional language at different levels of L2 proficiency. It is also worth mentioning in passing that, at the time, our prioritizing of the process of formulation represented somewhat of a novelty in the research on writing processes at large (be it L1 or L2 writing). We considered this to be an important gap to be filled because, as we argued, formulation is "the only compulsory activity while writing: Writers may decide to plan or not to plan, to revise or not to revise their texts, but there is no text at all if the writer does not attempt to transform ideas into language" (Manchón & Roca de Larios, 2007, p. 110). With the passing of time, the language component of composing has been made much more central in subsequent theorizing and model building (see, for instance, Hayes, 2012) as well as in empirical research agendas. After all, as noted by Arfé (2012), "a full understanding of writing, including its difficulties and disorders, must consider in detail the language processes and language mechanisms underlying the generation, formulation, and production of written text to communicate thoughts" (p. 573).

The various studies conducted within our program of research on writing processes provided us with a wealth of empirical data that distinctively showed the rich linguistic processing (and corresponding equally rich and most intriguing problem-solving behavior) that Cumming (1990) had anticipated as an integral component of composing. It was not until much later that we realized that this linguistic processing could be fruitfully reinterpreted from a language learning perspective. This new research orientation was initiated in Manchón and Roca de Larios's (2007) conceptual piece, where it was made explicit that the intention was to add to the "debate" initiated by other scholars. Thus, along the lines of the above mentioned modality-sensitive research agenda set up by Harklau (2002), together with Cumming's (1990) formulation of the language learning benefits of the linguo-cognitive activity that characterizes challenging writing, we attempted to elaborate on the psycholinguistic rationale for the anticipated language learning affordances of L2 writing in foreign language contexts, precisely the context in which we taught and conducted our research. We envisioned such learning potential to be closely linked to and to result from the theoretically predicted and empirically attested problem-solving nature of composing. We interpreted the learning potential of this problem-solving activity from prevalent SLA cognitive theoretical positions and, accordingly, we anticipated that L2 writers' attempt to transform ideas into

language could be "a process propitious to language development because of the learning mechanisms it would activate, and the corresponding possible changes in the L2 user's underlying linguistic system it might induce" (Manchón, 2017, p. 94). The resulting SLA-oriented, problem-solving-informed formulation of the language learning potential of L2 writing was elaborated more fully in later publications (Manchón & Williams, 2016; Roca de Larios, 2013). Importantly, however, the problem-solving approach to the study of writing as a site for language learning advocated in Manchón and Roca de Larios (2007) has only recently been adopted in empirical research, one example of which is the study reported in Chapter 10 (this volume. See also López-Serrano, Roca, & Manchón, 2019).

Manchón and Roca de Larios's (2007) position piece thus represented an additional contribution towards paving the way for subsequent theoretical and empirical developments. From the point of view of theory, this study attempted to provide the theoretical foundation for "the language learning potential of the problem-solving activity involved in frequent, repeated and guided practice in writing whole texts that form connected, contextualized, coherent, and appropriate pieces of communication" (p. 117). Additionally, we emphasized the mandate for the field to put theoretical predictions and tenets to the empirical test. Also from the perspective of empirical research, this study pointed to two new research avenues that, coincidentally, constitute at present areas of central interest in studies of writing and language learning, namely, task-related investigations together with research on written corrective feedback. Thus, in Manchón and Roca de Larios (2007) we interpreted previous work as pointing to the relevant role played by task-related variables, crucially including time-on-task, task design, and task implementation features. We also suggested the addition of a new element to the mix, namely, the provision of feedback as a key task variable in the domain of writing (as later discussed at length in Manchón, 2014). In effect, the study of written corrective feedback has gradually become a central area of research in the study of language learning affordances of L2 writing (see Bitchener, 2019; Bitchener & Storch, 2016, for overviews) and, by extension, the study of feedback through a language learning lens has become a central area of concern in feedback studies more generally (see Hyland & Hyland, 2019a, b).

Representative of these subsequent developments already brought to the surface in Manchón and Roca de Larios (2007), Chapter 5 offers a full elaboration of the rationales for the role of written corrective feedback in language learning, and Chapters 6 to 9 provide new empirical data on the role of task-related factors (e.g. task repetition – with and without access to WCF-, task modality, and task complexity) in bringing about language learning through writing.

Subsequent collective initiatives to drive theory and research forward

The three contributions previously referred to collectively constituted empirically-driven investigations of theoretical postulations for the role of writing (and, to a lesser extent, written corrective feedback) in L2 learning. Chronologically, they were followed by two professional initiatives that are regarded as instrumental in driving theory and research forward: I am referring to a book published in 2011, *Learning-to-write and writing-to-learn in an additional language* (Manchón, 2011a), and also to the special issue *Exploring L2 writing-SLA interfaces*, published in the *Journal of Second Language Writing* one year later (*JSLW*, 21, 2012).

These two collective publication projects served to advance conversations in several ways. To start with, they had a profound influence on strengthening fruitful interfaces between the fields of L2 writing and SLA research. The two collective projects had an equally profound influence on expanding research agendas in both fields by adding the study of writing to (mainly cognitively-oriented) SLA research agendas, as well as by adding the study of writing through a language learning lens to L2 writing research.

Regarding the strengthening of SLA-L2 writing interfaces, and reflecting shifts in professional debates, Ortega (2012), in her Epilogue to the *JSLW* special issue, started by noting that the "suggestion that an exploration of research interfaces between the fields of second language writing […] and second language acquisition […] is a worthwhile pursuit might raise eyebrows of incredulity among applied linguistic readers" (p. 404). In contrast, she concluded her contribution by observing that the *JSLW* 2012 special issue was "a telling sign that a new dialogue between the L2 writing and the SLA research communities is emerging, and one that is based on shared interests and genuine perceptions of mutual relevance that seemed unlikely just a few years ago" (p. 413). Ortega also optimistically anticipated that this cross-pollination would "eventually contribute to change in the landscape of both fields, by leaving a trail of valuable intellectual bridges among the relevant L2 writing and SLA research communities" (p. 413). Her predictions proved to be well grounded as worthy theoretical and empirical efforts have been put into building these "valuable intellectual bridges", including book-length treatments focused on both writing (e.g. Byrnes & Manchón, 2014b; Manchón, 2012) and written corrective feedback (Bitchener & Storch, 2016). The present book should be seen as an additional initiative intended to strengthen these SLA-L2 writing interfaces.

Regarding the expansion of research agendas by means of adding the study of writing as a site for L2 learning to SLA and L2 writing studies, the Editors of the *JSLW* explained the ultimate aim pursued with the 2012 special issue as follows (Manchón & Tardy, 2012, *v*):

> Fully aware of the interdisciplinary nature of the field of SLW, and also deeply convinced of the theoretical and practical relevance of exploring SLW-SLA interfaces, we asked a group of well- established scholars, with a sustained record of SLA-oriented SLW research, to jointly construct an account of the present and future work in the field by offering critical theoretical and methodological analyses of previous research, new empirical data, and directions for future research agendas

This same ultimate aim guided the edited collection *Learning-to-write and writing-to-learn in an additional language* (Manchón, 2011a). In his Preface to the book, Cumming (2012) observed that contributors "address the fundamental and intriguing paradox that L2 writing is not only an ability to acquire, teach, and assess – as is conventionally assumed- but L2 writing is also a means, context, and basis for learning, both of language and of writing" (pp. *ix–x*). He added that the theoretical deliberations and empirical analysis offered in the book "establish the groundwork and rationales to prepare new investigations into and to form new perspectives on the relationships between writing, language, and learning in diverse contexts and among varied populations around the world" (p. *xii*). Thus, from the point of view of theoretical developments, several contributions to these two publications (most notably Bitchener, 2012; Polio, 2012; Williams, 2012) provided a strong foundation for subsequent theoretical elaborations (e.g. Bitchener, 2019; Manchón, 2011b, 2014; Manchón & Williams, 2016). Manchón and Vasylets (2019), in their recent state-of-the art account, synthesize these theoretical developments along two dimensions. First, theorizing has revolved around the questions of (a) what is unique about and/or characteristic of writing and feedback that can lead to advancing language competences, and (b) what kind of learning can be expected to derive from written output practice and engagement with feedback, on the other (see Manchón & Williams, 2016 for a fuller analysis). Second, a key development that Manchón and Vasylets (2019) note refers to the recent initiatives to link theoretical predictions in the domain to SLA models, most notably Gass's (1997) cognitive model of input processing in the case of feedback (see Bitchener, 2019) and Leow's (2015) model of ISLA processes in the case of writing. Chapter 5 (this volume) adds to these previous initiatives with the first elaboration of a model of feedback processing within Leow's (2015) postulations.

From the perspective of empirical developments, research has grown in a variety of directions corresponding to the "new investigations [...] on the relationships between writing, language, and learning in diverse contexts and among varied populations around the world" that Cumming (2011, p. *xii*) considered to have spawned by the collective efforts in Manchón (2011a). These empirical developments are briefly synthesized in the next section as a way of offering the necessary background to situate the empirical studies in the current volume within current disciplinary

discussions. It might be worth reiterating that the Coda chapter (Chapter 17) assesses the book's contribution in terms of additions to past research and future developments that emerge from the theory and research discussed.

Empirical developments on writing as language learning: A synthetic overview

Figure 1 is a synthetic representation of the main empirical developments thus far (see Manchón and Vasylets, 2019, for a fuller analysis of the lines of research sketched out here). Two macro strand of research can be observed, namely those examining the act of writing itself, and those focused on the engagement with written corrective feedback (WCF).
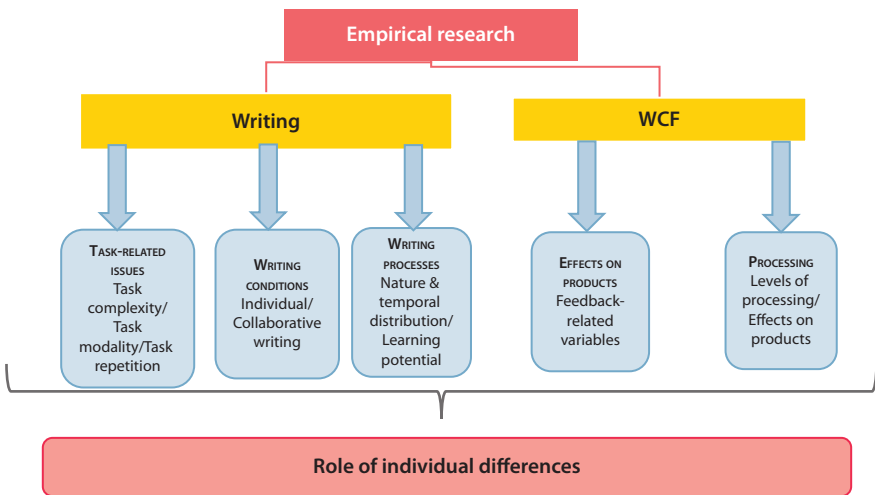
Figure 1. Overview of empirical research on writing and language learning

Regarding writing, two research preoccupations have attracted the most attention, identified as "task-related issues" and "writing conditions" respectively in Figure 1. As for the former, several studies have been conducted with the ultimate aim of ascertaining the learning outcomes of tasks (a) across modalities (e.g. Kormos, 2014; Tavakoli, 2014; Zalbidea, 2017), in some cases as moderated by task complexity factors (e.g. Vasylets, Gilabert & Manchón, 2017, 2019), as well as (b) across writing conditions, i.e. when writing individually or collaboratively, in both pen-and-paper and, gradually more often, digital environments (see Storch, 2013, 2016, 2018. See Stiefenhöfer & Michel, Chapter 11, and Saller, Chapter 12, this volume). More recently, as noted at several points in the above sections, a third strand has been added to research agendas, namely, a renewed interest in the study of writing processes (as

detailed in Chapter 14) with part of this research zooming into writing processes from the perspective of their learning affordances (López Serrano et al., 2019).

As depicted in Figure 1, the second main direction in empirical inquiries corresponds to the research interest in whether and, if so, how and why, WCF may lead to language development. This strand has grown exponentially and has resulted in an abundant body of empirical work (see reviews in Bitchener, 2019; Bitchener & Storch, 2016; Kan & Han, 2015; Liu & Brown, 2015) that clearly exceeds the number of studies on writing itself. The bulk of written feedback studies have looked into the (mostly short-term) effects of WCF on revised texts, while a growing number of empirical investigations have adopted a process orientation seeking to shed light on the engagement with and processing of the feedback received as well as on potential correlations between levels of processing and learning, operationalized in terms of the revisions undertaken (e.g. Caras, 2019; Cerezo et al., 2019; Coyle, Cánovas-Guirao, & Roca de Larios, 2018).

As shown in Figure 1, running across all the strands mentioned, we can identify an expanding body of empirical investigations that has responded to Kormos's initial (2012) call to make the study of IDs more central in L2 writing research intended to establish SLA-L2 writing interfaces. This research on individual differences has studied both writing and written corrective feedback (see Cho, 2018; Li & Roshan, 2019; Michel, Kormos, Brunfaut, & Ratajczak, 2019; Révész, Michel, & Li, 2017; Zabihi, 2018; Zalbidea, 2017, for working memory and writing/ feedback processing. Benson & DeKeyser, 2019; Kormos & Trebits, 2012; Sheen, 2007; Stefanou & Révész, 2015; Yan et al., 2019, for aptitude and writing/feedback appropriation. See also Ferris & Kurzer, 2019, for a review of studies addressing affective and attitudinal individual differences in feedback appropriation).

Collectively, the accumulated body of empirical work distinctively points to L2 writing as a favorable environment for language development from the dual perspective of the potential of writing itself (especially in comparison with the potential of speaking tasks, as reported in the available task-modality studies) and the learning benefits of learners' engagement with feedback (the research strand that has developed the most). The insights obtained thus far shed light mainly on short-term learning effects, especially effects on learning products – i.e. the characteristics of the texts written as a function of task-related variables –, or the nature of immediate revisions after receiving and processing/appropriating feedback as a function of feedback-related, task-related, and learner-related variables.

Yet, despite the abundant research on the connection between writing and language learning published in the last few years, many issues remain unresolved and many directions of research remain to be explored. Part of these needed developments relate mainly to the study of writing itself. Surprisingly, as noted in earlier sections, some of these future research directions on writing correspond precisely

to those signaled in or directly deriving from the pioneer contributions analyzed earlier on in the chapter. This applies most notably to, first, the partial neglect of those dimensions of writing itself purported to be conducive to language learning and that, in essence, correspond to, first, the intense linguo-cognitive activity that characterizes writing (as outlined in Cumming, 1990 and Manchón & Roca, 2007) and, second, to the needed educational, instructed SLA lens to be applied to the study of why and how writing may lead to language learning (as originally argued by Harklau, 2002). In what follows I outline the manner in which the current book seeks to make advancements in these domains.

Moving forward in research agendas on writing and language learning: The present book

This book seeks to make advancements in theory and research by directing the spotlight primarily on writing. Interestingly, the synthetic review of milestones provided in previous sections clearly shows that those scholars who pioneered the initiative to investigate the language learning potential of writing put forward claims and predictions that applied to writing itself and only marginally to feedback (mentioned only in Manchón & Roca de Larios, 2007). As shown in Figure 2, in pursuing its ultimate aim, the conception of the book takes stock of those main research avenues signaled in prior work, and attempts to advance professional discussions by (a) providing further theoretical and methodological reflections
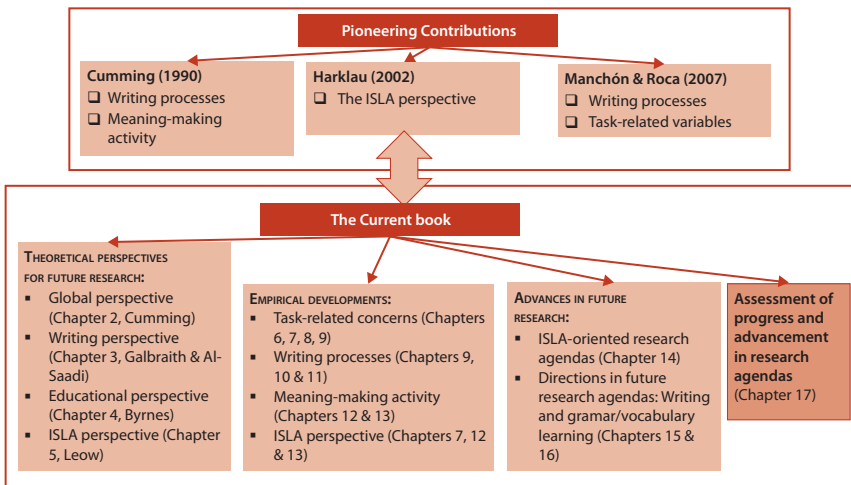

Figure 2. Overview of the current book in relation to key milestones in the study of writing and language learning

(the latter almost absent in extant research) that can help advance work in the domain, and (b) presenting cutting-edge empirical research intended to enhance our understandings of L2 learning through L2 writing. These three dimensions – theory, empirical work, and directions for future empirical inquiry – correspond to the central parts in which the book is divided, which are synthetically described below. The book also includes a final Coda chapter (Chapter 17) in which all the strings are pulled together in an attempt to (a) ascertain the way in which the theory and research reported in the book advances disciplinary discussions, and (b) point to avenues worth exploring in future research agendas.

Theory

Part I, *Advances in theoretical perspectives*, contains Chapters 2 through 5. Written by leading figures in L1 writing (David Galbraith and Zulaikha Al-Saadi), L2 writing (Alister Cumming, Heidi Byrnes), and instructed ISLA (Ron Leow), these four position pieces contribute theoretical reflections that can inform work in the domain. They do so in the following way:

– Alister Cumming (Chapter 2) provides an in-depth exploration of why writing may contribute to learning, in general, and language learning, in particular, and of the diverse lenses through which such connections can be approached. He reviews research and theories from the three macro perspectives (that clearly expand the theoretical perspective informing his pioneering 1990 study) and, on the basis of this analysis, he puts forward 10 tentative claims (scrutinized from a range of theories of learning) about diverse ways in which L2 writing may lead to L2 learning. He finishes with an innovative framework for an inquiry into the connection between L2 writing and L2 learning, which he encapsulates in the global dimensions of "learning *through*, *by*, *for*, or *with* writing" (p. 40). This framework, which represents a novel contribution to the field and to the book, serves to situate the theory and research contained in thevolume, as noted at several points in the rest of this chapter as well as in Coda chapter.

– David Galbraith and Zulaikha Al-Saadi (Chapter 3) contribute unique ideas about the qualities of knowledge involved in composing written texts, coupled with reflection on ways for research to address fundamental issues in L2 writing within the theoretical framework presented in their chapter. The distinction between episodic and semantic memories in composing written texts makes a unique contribution to the book, and represents a novel theoretical framework in which to situate the dimension of "learning by writing" identified by Cumming in the previous chapter. The suggestions put forward at the end of

the chapter about issues and designs for future work point to worthy future research avenues that are uncharted territory in studies of writing and learning in an additional language.

– Heidi Byrnes (Chapter 4) contributes another theoretical piece that, together with Leow's chapter, adopts an educational, curricular lens to the connection between writing and language learning. Taking an educational linguistics perspective, Byrnes articulates her proposal around the discussion of four areas that have the potential of illuminating possible immediate and long-term effects of L2 writing for L2 learning. All of them are, in one way or another, related to Cumming's original claims in his 1990 paper on the relevance of deepening our understanding of the meaning-making nature of writing as a central concern in understanding the connection writing and language learning. Byrnes's analysis leads to her proposal to inform future research in the four domains discussed in the chapter by a textually oriented theory of language, such as Systemic Functional Linguistics, and contemporary thinking in Complex Dynamic Systems Theory.

– Ron Leow (Chapter 5), in one of the few contributions to the book that reflect on both writing and feedback, focuses his analysis on another macro-dimensions identified by Cumming, i.e. "learning through writing", and, complementing Byrnes's chapter, situates the discussion in a curricular context. Following the pattern that is common to all the theoretical contributions in the volume, Leow offers a synthetic review of cognitive perspectives (including his own 2015 model) relevant for the analysis of the cognitive processes related to engagement with feedback, which is followed by a critical report of various dimensions of feedback studies premised on these theoretical underpinnings, and the proposal of a curriculum-, and process-oriented future research agenda. In this way, Leow's chapter serves to advance the process-oriented and classroom-based research directions signaled in the pioneering works by Cumming and Harklau.

Empirical developments

Part II, *Advances in empirical research*, includes Chapters 6 through 13. These contributions to the book, written by seasoned scholars and newcomers to the field, directly respond to the central theme of the volume and form a unified whole as they are all linked to the "learning through writing" and "learning by writing" dimensions identified by Cumming. They provide new empirical evidence on writing as a site for language learning obtained in studies framed in diverse theoretical perspectives (including cognitive SLA perspectives, Systemic Functional Linguistics, and models of L2 writing), conducted with diverse learner populations, in diverse

writing conditions and environments (including individual and collaborative writing in both pen-and-paper and digital environments), as well in diverse educational environments (CLIL settings, foreign language classrooms, university academic writing environments, and out-of-school settings).

Reproducing and adding to the contents of Figure 2 above, Figure 3 is a graphical representation of how the empirical studies in the current volume are situated in previous empirical work. As can be seen, four chapters (6, 7, 8 and 9) provide new empirical insights on task-related concerns (including task complexity, task modality, and task repetition), two chapters (11 and 12) study writing in collaborative writing conditions (although they do so from different theoretical perspectives), and three chapters (9, 10 and 11) add new insights on writing processes from a language learning angle. Additionally, Chapters 12 and 13 add a welcome new interest on the language learning potential of the meaning-making activity that is criterial to writing, hence expanding research on the "learning by writing" dimension identified by Cumming and elaborated in Byrne's contribution. More precisely:

–   Alberto Sánchez, Rosa M. Manchón, and Roger Gilabert (Chapter 6) examined the modality-dependency and proficiency-dependency of the learning affordances of task repetition in a study with secondary school and university students. The study findings (a) confirm and expand previous predictions regarding the modality-dependency of task repetition effects, and (b) distinctively point to enhanced language learning benefits of writing as compared
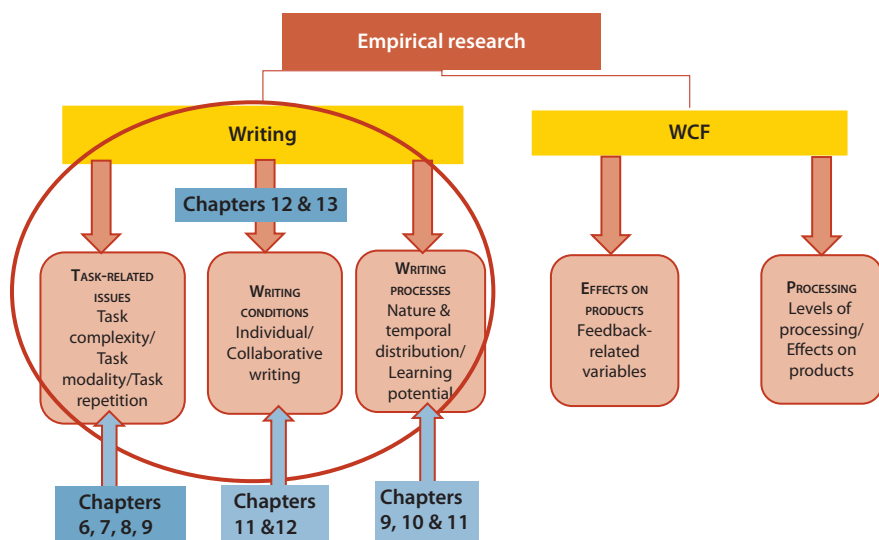

Figure 3. Framing of empirical studies in the current volume in previous research

to those created by speaking tasks. Additionally, this contribution serves to uncover the complex interaction of variables that appear to mediate task repetition learning outcomes in the written mode.

– Victoria Amelohina, Florentina Nicolás-Conesa, and Rosa M. Manchón (Chapter 7) also looked into task repetition, in this case longitudinally, in an out-of-school contexts, and analyzing the effects of 2 task repetition modalities (exact and procedural task repetition) aided with written corrective feedback (WCF). The longitudinal perspective adopted allowed the researchers to find differential effects of task repetition across time, differential effects on different dimensions of performance and, interestingly, differential appropriation of indirect WCF across time, which represents a novel insight in feedback studies. Methodologically, their study clearly points to the relevance of adopting a longitudinal, curricular perspective in the study of the connection between writing and language learning, a point made in various other contributions to the book (see especially Manchón & Leow, Chapter 14).

– Olena Vasylets, Roger Gilabert, and Rosa M. Manchón (Chapter 8) investigated the learning affordances of task modalities as mediated by task complexity. The researchers found commonalities across modalities (communicative success was found to be linked to lexical complexity and fluency across modes) as well as striking differences (adequacy was associated with accuracy only in speech and with propositional complexity only in writing). They also found that task complexity did not moderate the links between communicative adequacy and the CAF dimensions, which led them to conclude that mode of performance exerts more robust and potentially more predictable effects on L2 learners' performance than task complexity.

– Zalbidea (Chapter 9) also investigated task modality effects on L2 learners' processes and products as she examined task modality effects on (a) noticing of the target form, (b) target form incorporation, and (c) perceptions of task-induced demands in a grammar-focused task. Similar to the study by Vasylets et al. reported in Chapter 8, Zalbidea observed similarities and differences across modalities as she found that (a) both task modality conditions led to substantial noticing and form incorporation, along with similar ratings of task demands, and (b) participants in the writing condition were more accurate in incorporating the target form into their own output, hence providing further support to the greater language learning potential of writing tasks found in Sánchez et al.'s task repetition study reported in Chapter 6.

– Adding to Zalbidea's study of writing processes, and seeking to contribute to the research agenda originally put forward by Cumming (1990), Sonia López-Serrano, Julio Roca de Larios, and Rosa M. Manchón (Chapter 10) offer a detailed exploration of the mediating role of L2 proficiency on 21 EFL writers'

depth of processing and orientation of their strategic behavior while writing. Their analysis focuses on the language-related episodes they identified in the think-aloud protocols the participants produced during their individual completion of an argumentative writing task. The study results confirm previous predictions on the manner in which engaging in L2 writing may foster deep levels of language processing, as well as the proficiency-dependency of such processing.

– Laura Stiefenhöfer and Marije Michel (Chapter 11) contribute the third study on writing processes in the volume, in this case guided primarily by a methodological ultimate aim. Their study adds two crucial dimensions to those by Zalbidea and López Serrano et al. as they examined writing processes in digital, collaborative conditions, and triangulated data from text mining, interaction analyses, eye-tracking, and stimulated recall. They observed 8 international students in the UK using Google Docs for paired collaborative writing tasks. Their analytical procedure allowed them to uncover the richness and complexity of peer interaction and the nature of the emerging text in the writing conditions examined. The researchers interpreted their findings from the perspective of how data triangulation can facilitate the study of the learning affordances of collaborative, digital writing. In effect, the advantages of data triangulation for further research in the domain is brought up in the three studies of writing processes in the volume, an issue further discussed in the Coda chapter.

– Adding to Stiefenhöfer and Michel, Marcus Saller (Chapter 12) also examined the language learning affordances of collaborative L2 writing in the digital environment, although his study differs from the one reported in Chapter 11 in its aims, theoretical framing, longitudinal nature, and data sources used. Framed in Systemic Functional Linguistics, Saller's longitudinal, exploratory study sought to elucidate the language-learning affordances of collaborative (pair work) versus individual writing by advanced L2 university over one semester. Data consisted of the audio recordings of dyadic interaction, surveys, and expository essays. The results provide evidence of (a) the language learning potential of the complex meaning-making decisions and negotiations accompanied by deep problem-solving behavior that characterize collaborative writing, and (b) a marked differential development of syntactic complexity features characteristic of academic writing in the two writing conditions.

– A Systemic Functional Linguistics, longitudinal approach was also adopted in Rachel Whittaker and Anne McCabe's study reported in Chapter 13. The main differences with Saller's study relate to the population and context study (a secondary school CLIL context) and their focus on just individual writing. Whittaker and McCabe traced their participants' writing development during the 4 years of their compulsory education (on the basis of 64 texts by the same

16 students, on a topic from the history syllabus, collected yearly) by analyzing a key feature of disciplinary literacy, grammatical metaphor (GM). Based on the analysis of their rich data, the researchers interpreted their findings from theoretical postulates on writing as a site for language learning and conclude that the purported advantages shown for writing to learn language can well be exploited in contexts in which students learn content through an additional language, which they characterize as "a site for cognitively demanding writing tasks" (p. 327).

Future avenues

Part III, *Advances in Future Research Agendas*, includes 3 position papers, Chapters 14 to 16, written by seasoned scholars who have been key agents in the field. They each present a comprehensive research agenda for the future development of the domain. Manchón and Leow (Chapter 14) explicate why any principled inquiry into how writing may lead to language learning ought to be situated within an instructed SLA perspective, and what this positioning entails in terms of research methodological options to be taken and directions to be followed. This instructional perspective is further elaborated upon in the chapters by Schmitt and Polio, who delve into future research agendas on writing and vocabulary learning (Schmitt, Chapter 15), and writing and grammar learning (Polio, Chapter 16).

– Manchón and Leow (Chapter 14) argue for the relevance of framing L2 writing research associated with language learning as part of a language curriculum within an instructed second language acquisition (ISLA) perspective. In line with current ISLA theorizing, they argue that the field should prioritize the conduct of additional studies of the processing dimension of writing and, accordingly, propose several future avenues and methodological directions for a process-oriented agenda in the domain. With this, they link back to both the studies on writing processes in Part II and to Cumming's chapter in Part I as their claims and suggestions represent an elaboration of one of the global directions for research on the connection between writing and L2 learning identified by Cumming in Chapter 2, namely, the one focused on "attention, self-regulation, knowledge consolidation, or collaboration while composing", which involves "processing levels of attention, knowledge consolidation, and self-regulation" (p. 29).
– Diane Schmitt's contribution (Chapter 15) is guided by the ultimate aim of advancing a research agenda for investigating the vocabulary learning potential of writing instruction. She first reviews the challenges L2 writers face with vocabulary size, word knowledge, and lexical fluency from the perspective of

vocabulary research. The chapter then considers how vocabulary is commonly operationalized in writing studies and how vocabulary is treated in writing instruction. On the basis of these analyses, and following the same pattern as all chapters in Part III, Schmitt puts forward diverse proposals for a research agenda that aims to bring the concerns of vocabulary research and writing research into closer alignment.

–   Guided by the overarching question of "can writing facilitate the development of grammatical competence?", Charlene Polio (Chapter 16), after synthesizing disciplinary discussions of why writing should facilitate the acquisition of grammar, discusses six types of empirical studies that are related to this theoretical position. Similar to the approach followed by Leow and Manchón in Chapter 14, and Schmitt in Chapter 15, Polio then proposes a future research agenda along six research avenues offered as fruitful directions for new empirical investigations pursuing to shed a strong light on the facilitative effects of writing on grammatical development.

Closing commentary

As noted at the outset, the current book is guided by the dual ultimate aim of presenting cutting-edge empirical research, and providing theoretical and methodological reflections that can help advance research agendas. Therefore, the book is intended as a contribution to theory and research on writing and language learning and one of its distinctive features is that it combines in-depth theoretical and methodological reflections with new empirical findings in a single volume. In the Coda chapter (Chapter 17) I evaluate the contribution of the book in terms of what it adds to previous theoretical and empirical initiatives, and what directions for future inquiry it opens up. This Coda chapter will be the end of a journey that I now invite readers to inititate. It is hoped that through this journey readers interpret the kaleidoscope of orientations and methodologies in the empirical studies, together with the in-depth theoretical and methodological reflections included in the book, as evidence of the way in which our collective efforts have contributed to enhancing understandings of writing and language learning, on the one hand, and to strengthening ISLA/SLA-L2 writing interfaces, on the other.

References

Arfé, B. (2012). Looking into the text generation box to find the psycholinguistic (cognitive-language) writing processes. In V. W. Beninger, (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 573–578). New York, NY: Psychology Press.

Benson, S., & DeKeyser, R. (2019). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*, 23(6), 702–726. https://doi.org/10.1177/1362168818770921

Bitchener, J. (2012). A reflection on 'the language learning potential' of written CF. *Journal of Second Language Writing*, 21, 348–363. https://doi.org/10.1016/j.jslw.2012.09.006

Bitchener, J. (2019). The intersection between SLA and feedback research. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing. Contexts and issues* (pp. 85–105). Cambridge: CUP. https://doi.org/10.1017/9781108635547.007

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781783095056

Byrnes, H. (Ed.). (2013). Writing as meaning-making. Teaching to mean. Special Issue, *Journal of Second Language Writing*, 22(2).

Byrnes, H. (2014). Linking task and writing for language development: Evidence from a genre-based curricular approach. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 237–263). Amsterdam: John Benjamins.

Byrnes, H., & Manchón, R. M. (2014a). Task-based language learning – Insights from and for L2 writing: An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based L2 language learning – Insights from and for L2 writing* (pp. 1–23). Amsterdam: John Benjamins.

Byrnes, H., & Manchón, R. M. (Eds.) (2014b). *Task-based language learning- Insights from and for L2 writing*. Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.7

Caras, A. (2019). Written corrective feedback in compositions and the role of depth of processing. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 188–200). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-13

Cerezo, L., Manchón, R. M., & Nicolás-Conesa, F. (2019). What do learners notice while processing written corrective feedback? A look at depth of processing via written languaging. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 173–187). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-12

Cho, M. (2018). Task complexity and modality: Exploring learners' experience from the perspective of flow. *Modern Language Journal*, 102(1), 162–180. https://doi.org/10.1111/modl.12460

Coyle, Y., Cánovas-Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, 42, 25–43. https://doi.org/10.1016/j.jslw.2018.09.002

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, 39, 81–141. https://doi.org/10.1111/j.1467-1770.1989.tb00592.x

Cumming, A. (1990). Expertise in evaluating second language compositions. *Language Testing*, 7, 31–51. https://doi.org/10.1177/026553229000700104

Cumming, A. (2012). Preface. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. ix–xii). Amsterdam: John Benjamins.

Ferris, D., & Kurzer, K. (2019). Does error feedback help L2 writers? Latest evidence on the efficacy of written corrective feedback. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing. Contexts and issues* (pp. 106–124). Cambridge: CUP. https://doi.org/10.1017/9781108635547.008

Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11, 329–350. https://doi.org/10.1016/S1060-3743(02)00091-7

Hayes, J. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. https://doi.org/10.1177/0741088312451260

Hyland, K., & Hyland, F. (2019a). Contexts and issues in feedback on L2 writing. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing. Contexts and issues* (pp. 1–22). Cambridge: CUP. https://doi.org/10.1017/9781108635547.003

Hyland, K. & F. Hyland (Eds.) (2019b). *Feedback in second language writing. Contexts and issues*. Cambridge: CUP.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy. *Modern Language Journal*, 99(1), 1–18. https://doi.org/10.1111/modl.12189

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21(4), 390–403.

Kormos, J. (2014). Differences across modalities of performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes & R. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 193–217). Amsterdam: John Benjamins.

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62(2), 439–472. https://doi.org/10.1111/j.1467-9922.2012.00695.x

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge. https://doi.org/10.4324/9781315887074

Li, S., & Roshan, S. (2019). The associations between working memory and the effects of four different types of written corrective feedback. *Journal of Second Language Writing*, 45, 1–15. https://doi.org/10.1016/j.jslw.2019.03.003

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing*, 30, 66–81. https://doi.org/10.1016/j.jslw.2015.08.011

López-Serrano, S., Roca de Larios, J. & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically-motivated and empirically-based coding system. *Studies in Second Language Acquisition*, 41, 503–527.

Manchón, R. M. (Ed.) (2009). *Writing in foreign language contexts: Learning, teaching, and research*. Bristol, UK: Multilingual Matters.

Manchón, R. M. (2011a). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31

Manchón, R. M. (2011b). Writing to learn the language: Issues in theory and research. In R. M. Manchón (Ed.), *Learning-to-write and writing-to learn in an additional language* (pp. 51–82). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.07man

Manchón, R. M. (2012). *L2 writing development: Multiple perspectives*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781934078303

Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 27–52). Amsterdam: John Benjamins.

Manchón, R. M. (2017). The potential impact of multimodal composition on language learning. *Journal of Second Language Writing*, 38, 94–95. https://doi.org/10.1016/j.jslw.2017.10.008

Manchón, R. M., & Roca de Larios, J. (2007). Writing-to-learn in instructed language learning contexts. In E. A. Soler & M. P. S. Jordá (Eds.), *Intercultural language use and language learning* (pp. 101–121). Berlin: Springer. https://doi.org/10.1007/978-1-4020-5639-0_6

Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing: Implications for theory. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 102–129). Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859-008

Manchón, R. M., & Tardy, C. (2012). From the Editors. Exploring L2 writing-SLA interfaces. Special issue in the *Journal of Second Language Writing*, 21(4), *v*.

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. B. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge: CUP. https://doi.org/10.1017/9781108333603.015

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter. https://doi.org/10.1515/9781614511335-029

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing*, 45, 31–45. https://doi.org/10.1016/j.jslw.2019.03.002

Ortega, L. (2012). Epilogue: Exploring L2 writing–SLA interfaces. *Journal of Second Language Writing*, 21(4), 404–415. https://doi.org/10.1016/j.jslw.2012.09.002

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21, 375–389. https://doi.org/10.1016/j.jslw.2012.09.004

Révész, A., Michel, M., & Lee, M. (2017). Investigating IELTS Academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory. *IELTS Research Reports Online Series*, 44.

Roca de Larios, J. (2013). Second language writing as a psycholinguistic locus for L2 production and learning. *Journal of Second Language Writing*, 22(4), 444–445. https://doi.org/10.1016/j.jslw.2013.08.009

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Stefanou, C., & Révész, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *Modern Language Journal*, 99(2), 263–282. https://doi.org/10.1111/modl.12212

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847699954

Storch, N. (2016). Collaborative writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 387–406). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-021

Storch, N. (2018). Collaborative writing. In J. Liontas (Ed.), *The TESOL encyclopedia of English language teaching*. New York, NY: John Wiley & Sons.

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 217–236). Amsterdam: John Benjamins. https://doi.org/10.1002/9781118784235.eelt0395

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394–430. https://doi.org/10.1111/lang.12228

Vasylets, O., Gilabert, R., & Manchón, R. M. (2019). Differential contribution of oral and written modes to lexical, syntactic and propositional complexity in L2 performance in instructed contexts. *Instructed Second Language Acquisition*, 3(2), 206–227. https://doi.org/10.1558/isla.38289

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Yang, Y., Sun, Y., Chang, P., & Li, Y. (2019). Exploring the relationship between language aptitude, vocabulary size, and EFL graduate students' L2 writing performance. *TESOL Quarterly*, 53(3), 845–856. https://doi.org/10.1002/tesq.510

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication*, 35(1), 32–57. https://doi.org/10.1177/0741088317735836

Zalbidea, J. (2017). 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *Modern Language Journal*, 101(2), 335–352. https://doi.org/10.1111/modl.12389

Advances in theoretical perspectives

L2 writing and L2 learning

Transfer, self-regulation, and identities

Alister Cumming
University of Toronto

Can aspects of a second language (L2) be learned through writing the language? Research addressing this question has focused on (a) the transfer of knowledge and skills; (b) attention, self-regulation, knowledge consolidation, or collaboration while composing; or (c) development of identities within particular discourse communities and complex dynamic systems. This chapter reviews research and theories from these three perspectives, which suggest 10 tentative claims about diverse ways in which L2 writing may foster L2 learning. These claims are analyzed from the perspectives of behaviorist, cognitive, sociocultural, and complexity theories of learning.

The idea that writing in a second language can foster learning in that language has intrigued educators and researchers for several decades. The purpose of this chapter is to review the theories and research that have addressed this matter to date. Three perspectives have been established, focusing either on (a) the transfer of knowledge and skills; (b) attention, self-regulation, knowledge consolidation, or collaboration while composing; or (c) development of identities within particular discourse communities and complex dynamic systems. I suggest that each of these perspectives has multiple dimensions, ranging from (a) micro-levels of linguistic and cognitive resources to (b) processing levels of attention, knowledge consolidation, and self-regulation and on to (c) macro-levels of interactions with semiotic systems, other people, and identities within discourse communities. I discuss each of these perspectives and dimensions in sequence, leading to the formulation of 10 tentative claims about diverse ways in which L2 writing may foster L2 learning. I conclude by considering how four general theories of learning relate to these claims: behaviorist, cognitive, sociocultural, and complexity theories.

Transfer of knowledge and skills

A long-standing view on learning during L2 writing is the notion that transfer of lexico-grammatical structures from students' first languages accounts for many of the errors that frequently appear in students' L2 written texts. Ringbom (1987) notably articulated this view and demonstrated its realizations through detailed analyses of errors in the grammar and lexis of compositions written by Finnish students of English. The view follows from the Contrastive Analysis Hypothesis, initially advocated by Lado (1957) as a systematic method of analysis to explain the relative difficulties evident in learners' L2 production, providing a focus for teachers, students, and pedagogical materials to predict and act on such cross-linguistic "interference". Prescriptive applications of Contrastive Analysis to language education were later criticized and dismissed as reductive, lacking in appropriate research methods, empirically unable to determine difficulty in the L2, and neglecting consideration of intervening psychological and sociolinguistic factors (Schachter & Celce-Murcia, 1977; Wardaugh, 1970). Nonetheless, the central tenets of this view have been expanded and refined, for example, through Archibald's (1994) analyses of a broad range of discourse features in the English L2 texts of German L1 students, Kang's similar (2005) analyses of Korean L1 writers of English, and Ringbom's (2007) own more comprehensive perspective on the complexities of determining difficulty across different languages and writing systems. Kaplan (1966) famously applied the contrastive notion to the discourse level of rhetorical organization transferring from students' L1 to L2 (in ways that Belcher [2014] showed have now been largely dismissed for L2 writing).

A different, influential perspective on cross-linguistic transfer was articulated by Cummins' (1984) hypothesis that once students have developed cognitive-academic skills, such as written literacy, in one language they remain available as a basis for performing such skills in additional languages. Cummins was concerned about evidence that many young minority-language children who do not develop literacy and academic skills in their home languages struggle and take an inordinately long time to develop these abilities in unfamiliar majority-language contexts at school, particularly when teachers may not be aware of such limitations among students who have acquired more readily accessible basic, interpersonal conversational abilities in the majority language. Cummins has therefore advocated the importance of literacy development in children's first or home languages as a foundation for their academic success in a second language. Cummins' "linguistic interdependence hypothesis" has been applied to various studies of children's bilingual writing (Buckwalter & Lo, 2002; Hornberger, 2003; Pérez, 2004; Verhoeven, 1994) as well as adults' L2 writing (Berman, 1994; Cumming, 1989, 1990, 2001; Hall, 1990; Whalen & Menard, 1995), demonstrating that people who have developed composing

processes and literacy in one language readily apply them while writing in a second language even in spite of lacking relevant lexical and grammatical proficiency.

Indeed, "interdependence" may be a more appropriate word than "transfer" is, given that recent neurolinguistic studies indicate that the neural networks in bilinguals' brains are activated in both of their languages concurrently and spontaneously while processing challenging semantic and lexical tasks (Dong, Gui, & MacWhinney, 2005; van Heuven & Dijkstra, 2010; Kroll & Bialystok, 2013; Linck, Hoshino, & Kroll, 2008; Strijkers, 2016). Further, as Larsen-Freeman (2013) has argued, "transformation" may be a more appropriate concept to explain people's adaptive applications of new knowledge in different contexts than is the transportation metaphor of exporting implied by the term "transfer". L1 and L2 knowledge and skills interact, overlap, and develop across contexts depending on people's experiences, genres of interaction, and identities adopted. For studies of L2 writing, Kobayashi and Rinnert (2012) have showed that the influences of L1 and L2 writing abilities can go in both directions for Japanese writers of English in the sense that abilities acquired in the L2 can affect writing in the L1, and vice-versa. Likewise, Sasaki (2011) showed that Japanese learners of English who have studied English abroad develop qualitatively different motivations for writing in English, senses of their intended audiences while writing in English, and views on their own identities as users of the L2 compared to similar students who have not traveled outside of Japan. More generally, the importance of examining how students transfer their learning of L2 writing from instructional contexts to academic courses and workplace roles has been explored in numerous studies by James (2010, 2014) and others, addressing a fundamental justification for the teaching of L2 writing as well as to recognize its value and limitations. L2 writing is not only a skill to learn but also a means of learning the language and acquiring and conveying relevant content in academic as well as workplace settings (Hirvela, 2011; Manchón, 2011).

L2 learning while composing

Most research on the learning of an L2 while writing has focused on what learners do, attend to, and think about while they compose written texts. One line of inquiry has observed that humans have limited attentional capacities when performing complex tasks, so when writing in a second language, learners' attention is inevitably and frequently devoted to addressing their limitations in lexical and grammatical resources, trying (perhaps with difficulty) to find or spell the right words or decide on appropriate grammatical forms. Whalen and Menard (1995) produced early evidence to show that during L2 writing students' attention is drawn most frequently to lexical and grammatical aspects of their emerging texts rather

than attending to discourse or rhetorical concerns, compared to the same students' writing in their L1. Fitzgerald (2006) reviewed considerable research that came to similar conclusions with school-age writers of L2 English. The premises of this view are similar to those that Robinson (e.g., 2011) has investigated in his formulation of the Cognition Hypothesis about task performances in second languages. Accordingly, as Robinson has argued for second language acquisition generally, a more nuanced and multi-faceted view of attentional resources and task complexity has to be adopted rather than simply assuming that L2 writers are inordinately consumed or distracted by their deficits in linguistic knowledge. Learners may make an L2 writing task more (or less) challenging than it needs to be and so learn about the language and develop composing skills in the process (Uzawa & Cumming, 1989).

The predominant view in recent research about L2 writing and learning is that the conditions and processes of composing present unique circumstances for L2 learning (see chapters 1 and 17, this volume). The permanency, self-controlled pace, and expectations for precision of expression in writing can prompt learners to attend explicitly to language forms and meanings together to refine and consolidate their L2 knowledge and skills (Williams, 2012). Particularly conspicuous opportunities for learning the L2 appear when writers evaluate forms of the L2 in relation to their intended meanings, search earnestly to find the best words to express ideas, and switch purposefully between languages to make principled decisions (Cumming, 1989, 2001, 2013; Knutson, 2006; Murphy & Roca de Larios, 2010; Qi, 1998; Wang, 2003; Wang & Wen, 2002). Swain (2006) and colleagues (e.g., Swain & Lapkin, 1995) have called this learning potential "languaging" or "comprehensible output", building on the Vygotskian principle that learners' own thinking about their language production mediates their learning of the language – either as private or inner speech when composing alone (cf. de Guerrero, 2018) or as collaborative dialogue when composing together with others (cf. Storch, 2013).

Viewed from theories of usage-based learning and complex dynamic systems, repeated iterations of thinking episodes while composing in the L2 – that for example, involve searching for the best words to express ideas, extending one's linguistic resources, and producing comprehensibly accurate forms and discourse – could lead learners to restructure interacting aspects of their language and semiotic systems progressively over time (Baba & Nitta, 2014; Cumming, 1990; Fogal, 2017; Verspoor, Schmid, & Xu, 2012). Repeated, engaged experiences in writing may also prompt learners to develop increasingly complex mental models of L2 writing (Nicolás-Conesa, Roca de Larios, & Coyle, 2014) and exercise agency to shape personally significant innovations in their writing styles, the genres they produce, and the contexts in which they communicate (Tardy, 2016). Further dimensions and extensions of the self-regulating potential for L2 learning are evident in research that has demonstrated that while composing L2 writers can restructure their language

and ideas (Roca de Larios, Murphy, & Manchón 1999), establish and act productively on personal goals for learning (Cumming, 2006), and learn from exemplary text models and peer feedback (To & Carless, 2015).

Ellis (2019) elucidated how usage-based theories of language learning feature principles of "embodiment, environmental embeddedness, enaction, social enculturation, situatedness, and distributed cognition" to explain "how we learn language while engaging in communication" (p. 49). Such learning occurs as people form rich, complex representations in their memories through language experiences that are engaging, frequent, imageable, contextualized, goal-directed, and involve interactions with others (p. 48). For example, extensive, repeated, and engaged experiences of reading and writing argumentative texts may prompt the acquisition of that genre: This potential is evident in Beigman Klebnev, Ramineni, Kaufer, Yeoh and Ishizaki's (2019) natural language processing analysis that revealed fundamentally similar rhetorical functions for argumentative writing across samples of hundreds of compositions written for formal language tests, graduate students' course papers in a range of disciplines, and opinion editorials in the *New York Times*.

Identities in discourse communities

In addition to the psycholinguistic dimensions of L2 learning reviewed thus far, L2 writing and learning are also social phenomena. People develop abilities to write in second languages over relatively long periods of time within and for specific societal contexts and particular academic or job-related purposes. One aspect of such development is that of novices gaining and establishing membership into a discourse community by progressively adopting its specific terminologies, genre and register conventions, and discourse practices in their writing. In-depth studies by Leki (2007), Macqueen (2012), and Tardy (2009) have documented over the period of years how several L2 learners progressively acquired and used in varied ways personal repertoires of formal lexical phrases and genres while engaging in writing tasks in English for their academic courses at university. This productive, usage-based process of L2 learning while composing has made me wonder if the more aberrant behavior of patchwriting should also be viewed as a process of language learning while writing, involving the (more or less) indiscriminate borrowing and reformulation of lexical phrases from source texts (Cumming et al., 2018; Flowerdew & Li, 2007; Li & Casanave, 2012; Mizumoto, Hamatani, & Imao, 2017; Shi, 2010).

Writing competently in a second language is perhaps the chief way in which scholars, professionals, or technicians can signal their identities as members of a specialized discourse community. Producing such L2 writing, however, involves

the handling and transformation of relevant content knowledge, ideas, and genres as well as taking on new senses of personal identities. Such processes of writing and knowledge development feature vividly in longitudinal case studies of writing development in specific academic fields (e.g., Kibler, 2014; Leki, 2007; McCarthy Young, & Leinhardt, 1998; Tardy, 2009). Ivanic (1998) documented how a few undergraduate students struggled while performing written assignments over several years to establish new dimensions of their identities in respect to what she termed to be their (a) autobiographical selves, (b) senses of themselves as authors, and (c) adoptions of appropriate discourse conventions. Gentil (2005) likewise traced the intricate, consequential decisions that bilingual writers must make to develop advanced writing skills and become members of alternative linguistic communities for academic purposes in one or both of their languages. Harklau (2002) outlined ways in which writing should be considered – in conjunction with but also above and beyond oral interactions – for young English language learners in American classrooms in terms of variation between learners' and target language norms, multimodal communications, language socialization, and interactionist concepts of learning. In almost any social or pedagogical context that involves writing, writing practices are modeled, supported, and negotiated with others in ways that can facilitate language learning. Socio-cultural theories describe such learning as the gradual internalization of inter-psychological processes (socially) into a person's independent, intra-psychological abilities. An especially distinctive account of such language learning through writing appeared in Parks and Maguire's (1999) analysis of a newly hired francophone nurse learning to write nursing notes in English on the job in a hospital in Montreal. The nurse modeled the genre of his writing, including uses of specialized terminology and formats, on notes produced by more experienced peers as his emergent L2 writing was coached and corrected regularly by others working in the hospital. In educational practices, numerous analyses of one-on-one tutoring have documented the modeling, verbal supports, routine practice, and gradual fading away by tutors that facilitate learners' writing improvement and independent responsibilities (Aljaafreh & Lantolf, 1994; Cumming, 2012; Jun, Ramírez, & Cumming, 2010). Storch (2013) consolidated her many previous research studies to show how collaborative writing among peers in language classrooms presents optimal opportunities through discussion, deliberation, and cooperation for learning an L2 as well writing skills. Li and Zhu (2017) have demonstrated how such learning opportunities emerge when writing in wikis, and Cho (2017) has elucidated how collaborative learning and writing processes materialize in other multimedia contexts for academic writing outside of classrooms.

Ten claims and relevant theories of learning

The issues emerging from the foregoing review lead to ten claims that might be posed as tentative interpretations about the diverse ways in which writing in a second language may facilitate learning of that language. Although the substance of these interpretations have already been put forward in the publications cited, their realizations, precise nature, and value require future inquiry. The ten claims could, for example, be formulated as hypotheses and then evaluated empirically, but to do so, I believe the claims need to be articulated more fully in reference to established theories of learning, as I will start to do to conclude this chapter.

To summarize the points above, prior studies have suggested that L2 learning may occur through L2 writing,

- As processes of transfer:
    1. Similarities or differences between L1 and L2 explain frequent errors (as "interference") in L2 writing.
    2. Cognitive-academic literacy abilities established in L1 transfer to L2.
    3. L1 and L2 knowledge and skills interact, overlap, and can transfer bidirectiionally across contexts and languages depending on experiences, genres, and identities.
- Through processes of composing:
    4. People have limited cognitive capacities, constraining L2 writers' attention to text-level features while composing – or prompting them to upgrade their thinking and effort.
    5. The permanency, self-controlled pace, and expectations for precision of expression in writing can prompt learners to attend explicitly to language forms and meanings together to refine and consolidate their L2 knowledge and skills.
    6. While composing, L2 writers can restructure their language and ideas, act on personal goals for learning, or learn from exemplary text models or peer feedback.
    7. People restructure complex, dynamic, interacting language, discourse, and semiotic systems through repeated usage and personal agency in communicative interactions.
- And through people forming identities as members of discourse communities:
    8. Novices gaining membership into a discourse community progressively adopt its terminologies, registers, and discourse practices in their writing.
    9. Writing practices are modeled, supported, collaborated, and negotiated with others whereby inter-psychological processes are gradually internalized into independent, functioning abilities.

10. Writing to produce academic or professional knowledge involves transforming that knowledge to signal group membership through new senses of self, discourse, and authorial functions relevant to social contexts.

What theories might support these claims? Over past decades, scholarly discussions of learning have been dominated by four general theories, each of which relates to studies of L2 learning through L2 writing: behaviorist, cognitive, sociocultural, and complexity theories.

Behaviorist theories of learning

Behaviorist learning theories were initially claimed as support for contrastive analyses (e.g., Lado, 1957), but applied linguists came to question these foundations, particularly their applications to pedagogy through error analysis (Schachter & Celce-Murcia, 1977; Wardaugh, 1970). Errors in L2 speech or writing are difficult to identify with precision, ascribe for causes (e.g., as sources of L1 transfer or difficulty), explain comprehensively to learners (without extensive metalinguistic terminology, though that can be taught, cf. Schleppegrell, 2016), address without negative reinforcement or distraction from other learning opportunities, and remedy through practice (Truscott, 1996). Despite these concerns, great interest in corrective feedback on L2 writing has continued – as an aspect of instruction with the potential to promote learning of the language, albeit indirectly, by raising learners' awareness about errors or prompting revisions of a text after a teacher's or peer's feedback on their writing. Research on corrective feedback of writing has been voluminous (Bitchener & Ferris, 2012; Ferris, 2012; Lee, 2013; Leki, Cumming, & Silva, 2008), producing evidence of "gains in writing development" and that "focus on form and content is more effective than an exclusive focus on form" (Biber, Nekrasova, & Horn, 2011, i; see also Kang & Han, 2015).

But studies of corrective feedback have not followed behaviorist theories of learning. The central tenet of behaviorism is that learning occurs through operant conditioning – either as positive reinforcement (e.g., praise or repetition) or negative reinforcement (e.g., punishment or ignoring) (Skinner, 1957). Teachers' responses to students' writing may well involve praise as positive reinforcement (though that would hardly be called "corrective"). But corrective feedback of writing is seldom conceived as punishment (except perhaps by students who dislike or do not appreciate it). Moreover, teachers' providing corrective feedback is obviously contrary to the idea of ignoring aberrant behaviors because such feedback explicitly draws attention to errors. In short, behaviorist theories of learning have no particular value to explain or guide ideas about L2 learning through L2 writing. Instead,

research on corrective feedback and studies of cross-language transfer have tended to follow cognitive theories of learning, as has most other inquiry into L2 learning through L2 writing.

Cognitive theories of learning

Various cognitive theories of learning exist but are perhaps most fully exemplified in Anderson's theory of skill learning, called Adaptive Control of Thought-Rational (Anderson, 1982, 1995; Anderson, Bothell, Byrne, Douglass, Labiere, & Qin, 2004). Key, relevant tenets of this and other cognitive theories of learning are that declarative knowledge and procedural skills involve processing information; working memory limits the focus of attentional resources, the processing of information, and goal-directed behaviors contextually or ecologically; new behaviors build on existing knowledge and skills; extensive practice and experiences produce fluent, automatic behaviors; and problems in performance can be analyzed through self-regulation or other-regulation, for example, via heuristics that apply declarative understanding as cognitive restructuring to improve procedural behaviors. See Cumming (2016) and MacArthur and Graham (2016) for recent reviews related to L2 writing and L1 writing, respectively.

From a cognitive science perspective, transfer of learning is not reusing information but rather people's adaptive transformation of existing abilities to new experiences and environments (De Palma & Ringer, 2011; James, 2014; Larsen-Freeman, 2013). Thus, people apply writing abilities they have already established to new languages, tasks, contexts, or purposes (Cummins, 1984). As Kroll and Bialystok (2013, p. 1) proclaimed in reviewing considerable research on bilingual neural processing, "bilinguals activate information about both languages when using one language alone." When writing, if lacking appropriate words or language forms in an L2 or L1, people can pause to regulate their writing performance, search their memories for relevant resources, apply heuristic search strategies across first or second languages, and restructure tentative verbal formulations to identify and resolve perceived problems and so confirm or extend their knowledge (Chenoweth & Hayes, 2001; Cumming, 1989, 1990; Roca de Larios, Murphy, & Manchón, 1999). Such episodes of self-control, reflection, and cognitive restructuring provide potential opportunities for learning the L2 while writing it. More broadly, the learning potential is shaped by a person's motivation, goals for writing and learning, assistance from others (including teachers, peers, or collaborators), and imagined expectations from readers (Cumming, 2006; Hayes, 2012). These elements are further facilitated during stages of planning, information-gathering, drafting, and revising writing. Practice writing particular genres of writing, having relevant knowledge

about a topic (gained from personal experience, searching sources of information, or other people), and developing a sense of identity as a member of a discourse community that values that information and writing about it further enhance the writing and abilities to learn from doing it (as various L1 studies have showed: Graham & Herbert, 2011; Klein, Arcon, & Baker, 2016; Newell, Beach, Smith, & VanDerHeide, 2011; McCarthy, Young & Leinhardt, 1998).

Sociocultural theories of learning

Whereas cognitive science has tended to conceive of learning as individual information processing, self-regulation, and problem solving, sociocultural theories of learning situate learning primarily within cultural contexts and interactions. People learn through supportive interactions with others, mediated by material and symbolic tools and concepts, leading to the internalization of inter-psychological processes and knowledge (Lantolf, 2000; Vygotsky, 1978; Wertsch, 1985). Learning occurs as mediation within a person's unique zone of proximal development, scaffolded by others' support (e.g., teachers or collaborations in pairs or writing groups, Storch, 2013) and by oneself (through inner speech, de Guererro, 2018, or collaborative dialogue, Swain, 2006). Developing writing abilities involves progressive socialization into culturally appropriate ways of interacting with texts, people, social contexts, and knowledge (Bazerman, 2016; Duff, 2010; Prior, 2006). More broadly, writing occurs within activity systems that have particular rules of conduct, community memberships, and division of labor that also mediate writing and learning, whether in contexts of classrooms or of workplaces, and require experience, apprenticeship, or membership to participate in (Engeström, 2008, 2015; Haneda, 2007; Lei, 2008; Park & De Costa, 2015). Moreover, writing occurs in socially established and recognizable genres, which require the consolidation of relevant rhetorical, formal, subject-matter, and process knowledge to acquire and perform competently in schools (Christie, 2012; Schleppegrell, 1994) and higher education (Tardy, 2009, 2016).

Complexity theories of learning

Languages, writing, and learning are complex, varied, and ever-changing phenomena. Biliteracy phenomena are all the more so, as Hornberger and colleagues (2003) have explicated, involving dynamic, interacting systems of: identities, power, positions, self-organization, writing systems, communication modalities, maturation,

and potential for transformation; relationships and connectedness with others in local, extended, and historical social networks; and institutions, social classes, and societies. Complexity theories remind us that learning an L2 through writing could never be predictable or uniform given the variety of situations, populations, societies, forms of writing, and combinations of first and second languages that exist around the world and have existed in the past or will in the future as new technologies emerge, proliferate, and combine with other media of communication and interaction (Hornberger, 2003; Larsen-Freeman & Cameron, 2008).

One way of approaching this complexity is through post-structural considerations of the multiple identities that people adopt, invest in, or resist in relation to the exercise of power and privilege in institutions and society as well as multimodal and globalized chains of language, personal positions, and social affiliations (Norton & Toohey, 2011). This perspective invites a focus on L2 writing and learning among marginalized populations, such as immigrant women (Cumming & Gill, 1991; Norton, 2013) or economically disadvantaged youth (Cumming, 2016; De Costa, 2010; Dressman, Wilder, & Connor, 2005; Stein, 2008). Other viable points of investigation have been intersections between personal and societal commitments to vernacular and global interests in literacy (Canagarajah, 2004; Janks, 2010) or uses of first and second languages in workplace genres and communication systems (Parks & Maguire, 1999; Winsor, 2003).

A different approach is to understand the complexity of L2 learning through writing in relation to comprehensive theoretical frameworks that span psycholinguistic, sociolinguistic, and temporal perspectives. Hornberger's (1989, 2003) continua of biliteracy have proved to be a useful analytic lens for many insightful studies, distinguishing between multiple layers, patterns, and variations of bilingual contexts, individual development, and communication media around the world. Hornberger has shown how social contexts for biliteracy have differing micro, individual and macro, societal dimensions; demarcate and combine oral and literate uses of languages; and involve specialized situations, relationships, and functions for monolingualism or bilingualism. Individuals' development of abilities to read and write in more than one language also vary according to the relative emphasis that education and experiences place on a first or second or additional languages and on reading, oral, or written production of those languages, and whether the languages are acquired simultaneously or successively at different points in the lifespan, involve different or similar scripts, and have convergent or divergent linguistic and rhetorical structures. Another comprehensive framework is Bronfenbrenner's (1979) ecological systems theory, which Wilson (2013) adopted to study immigrant adolescents learning literacy in their schools, homes, and social networks. Bronfenbrenner's (1979) theory analyzes: Microsystems (family, friends,

neighbourhoods) and their interconnections in more extensive mesosystems, exosystems involving others with whom a person does not have direct contact or control, and macrosystems of institutions, social class, and societies; as well as chronosystems of changes over one's lifespan and over history.

Summary thoughts

If I can be allowed a little playing with prepositions, I suggest that L2 learning while writing can be considered either narrowly or broadly, that is, as learning *through*, *by*, *for*, or *with* writing. Narrowly, L2 learning *through* writing appears to happen *through* cognitive problem solving and restructuring, applications and enhancements of self regulation, and collaborations with others while writing. Viewed broadly, L2 learning can be said to involve learning *by* writing (from a usage-based, activity theory, or identities perspective), *for* writing (from a perspective of motivation, purpose, or identities), and *with* writing (in complex, dynamic systems).

Caution needs to expressed, though: Despite the many proposed perspectives on how L2 writing may relate to and prompt L2 learning, causality remains unproven. Consensus from syntheses of the abundant research that has accumulated about the potential effects of L1 writing on conceptual learning in academic contexts is that researchers have come up short on trying to prove that L1 writing actually causes learning of new knowledge. A meta-analysis of that research by Bangert-Drowns, Hurley, and Wilkinson (2004) concluded that the instructional benefits of writing to learn for L1 academic purposes are not causal per se but rather center on the creation of contexts that promote thoughtful reflection about, and articulation of ideas through, writing about specific topics. More recent syntheses of research have focused on evaluating the quality and conditions of effective writing practices that aim to promote conceptual learning. Specifically, these syntheses have evaluated the presence of three criteria: interactive writing processes (i.e., communicating purposefully with other people), meaning-making (requiring original or critical thought), and clear expectations (so learners know criteria to direct their learning and evaluation) (Anderson, Anson, Gonyea, & Payne, 2015). Gere, Limlamai, Wilson, Saylor, and Pugh (2018) established that these components of writing "correlate highly with the greatest learning gains among students" (p. 31), but their analyses of prior studies questioned whether writing to learn could be equally effective for all learner populations, abilities, and situations. They also observed that the operationalization of constructs of learning or knowledge in pre-post, control-group research designs tend to involve conflicting measurement methods such as matching (a) shallow, closed-response or recall items about ideas

or topics against (b) considerably deeper and individually variable processes of thinking while writing. Further methodological and conceptual refinements along these lines, together with others outlined throughout this chapter and book, might lead researchers to produce evidence about the benefits of L2 writing for L2 learning. But that conclusion still remains to be proved firmly.

Numerous issues warrant further attention, as the remaining chapters in this book demonstrate. One issue is whether the potential for L2 learning through writing is qualitatively different at different levels of L2 proficiency or even whether it might be quantitatively different in the sense that some kinds of learning described above require a certain level of L2 writing fluency and foundation lexical resources, as Cummins' (1984) "threshold hypothesis" proposed. Similarly, are certain ages of maturity, levels of education, or literate abilities required to facilitate L2 learning through L2 writing? Further, might there be a point at high levels of L2 proficiency when such processes exhaust their usefulness? A second issue is that metacognition has been emphasized in most analyses and explanations, particularly those involving self-reports about composing. Could research aim to address more basic, implicit or tacit aspects of cognition such as attention, neural processing, or self-regulation through, for example, brain imaging, eye-tracking, or response-time methods? Alternatively, could distinctly qualitative methods of inquiry, such as life-histories or ethnographies of classrooms or multilingual workplaces, illuminate L1/L2 interactions involving significant learning at key incidents in the lifespan and within specific discourse communities? Training studies could also be revealing, for example, to address strategic goals, specific heuristics, or editing of texts for L2 learning while writing. A final point is that variations in cross-linguistic combinations have scarcely been explored, either through within-subjects designs involving the same people composing similar tasks in first and second languages, or through comparisons between L1s and L2s that are either highly similar or different in their scripts and discourse structures. Likewise, variations in modes of writing need to be evaluated, not only across contrasting genres of texts or conditions of composing but also in new and varied multimedia communications.

References

Aljaafreh, A., & Lantolf, J. (1994). Negative feedback as regulation and second language learning in the Zone of Proximal Development. *Modern Language Journal*, 94, 465–48. https://doi.org/10.1111/j.1540-4781.1994.tb02064.x

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406. https://doi.org/10.1037/0033-295X.89.4.369

Anderson, J. R. (1995). *Learning and memory*. New York, NY: John Wiley & Sons.

Anderson, J. R., Bothell, D., Byrne, M., Douglass, S., Labiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036–1060. https://doi.org/10.1037/0033-295X.111.4.1036

Anderson, P., Anson, C., Gonyea, R., & Payne, C. (2015). The contributions of writing to learning and development: Results from a large-scale multi-instrumental study. *Research in the Teaching of English*, 50, 199–235.

Archibald, A. (1994). *The acquisition of discourse proficiency: A study of the ability of German school students to produce written texts in English as a foreign language*. Frankfurt: Peter Lang.

Baba, K., & Nitta, N. (2014). Phase transitions in development of writing fluency from a complex dynamic systems perspective. *Language Learning*, 64, 1–35. https://doi.org/10.1111/lang.12033

Bangert-Drowns, R., Hurley, M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, 74, 29–58. https://doi.org/10.3102/00346543074001029

Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23). New York, NY: Guilford Press.

Beigman Klebnev, B., Ramineni, C., Kaufer, D., Yeoh, P., & Ishizaki, S. (2019). Advancing the validity argument for standardized writing tests using quantitative rhetorical analysis. *Language Testing*, 36, 125–144. https://doi.org/10.1177/0265532217740752

Belcher, D. (2014). What we need and don't need intercultural rhetoric for: A retrospective and prospective look at an evolving research area. *Journal of Second Language Writing*, 25, 59–67. https://doi.org/10.1016/j.jslw.2014.06.003

Berman, R. (1994). Learners' transfer of writing skill between languages. *TESL Canada Journal*, 12, 29–46. https://doi.org/10.18806/tesl.v12i1.642

Biber, D., Nekrasova, T., & Horn, B. (2011). *The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis*. TOEFL iBT RR-11–05. Princeton, NJ: Educational Testing Service.

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. New York, NY: Routledge. https://doi.org/10.4324/9780203832400

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Buckwalter, J. K., & Lo, Y. G. (2002). Emergent biliteracy in Chinese and English. *Journal of Second Language Writing*, 11, 269–293. https://doi.org/10.1016/S1060-3743(02)00088-7

Canagarajah, S. (Ed.). (2004). *Reclaiming the local in language policy and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chenoweth, A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18, 80–90. https://doi.org/10.1177/0741088301018001004

Cho, H. (2017). Synchronous web-based collaborative writing: Factors mediating interaction among second-language writers. *Journal of Second Language Writing*, 36, 37–51. https://doi.org/10.1016/j.jslw.2017.05.013

Christie, F. (2012). *Language education through the school years: A functional perspective. Language Learning*, 62 (Supplement 1).

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, 39, 81–141. https://doi.org/10.1111/j.1467-1770.1989.tb00592.x

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511. https://doi.org/10.1177/0741088390007004003

Cumming, A. (2001). Learning to write in a second language: Two decades of research. *International Journal of English Studies*, 1, 1–23.

Cumming, A. (Ed.). (2006). *Goals for academic writing: ESL students and their instructors*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.15

Cumming, A. (Ed.). (2012). *Adolescent literacies in a multicultural context*. New York, NY: Routledge. https://doi.org/10.4324/9780203120033

Cumming, A. (2013). Multiple dimensions of academic language and literacy development. In L. Ortega, A. Cumming, & N. Ellis (Eds.), *Agendas for language learning research. Language Learning* 63(Supplement 1), 130–152.

Cumming, A. (2016). Theoretical orientations to L2 writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 65–88). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-006

Cumming, A., & Gill, J. (1991). Learning ESL literacy among Indo-Canadian women. *Language, Culture and Curriculum*, 4, 181–200. https://doi.org/10.1080/07908319109525104

Cumming, A., Yang, L., Qiu, C., Zhang, L., Ji, X., Wang, J., Wang, Y., Zhan, J., Zhang, F., Xu, C., Cao, R., Yu, L., Chu, M., Liu, M., Cao, M., & Lai, C. (2018). Students' practices and abilities for writing from sources in English at universities in China. *Journal of Second Language Writing*, 39, 1–15. https://doi.org/10.1016/j.jslw.2017.11.001

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, UK: Multilingual Matters.

De Costa, P. (2010). From refugee to transformer: A Bourdieusian take on a Hmong learners' trajectory. *TESOL Quarterly*, 44, 517–541. https://doi.org/10.5054/tq.2010.226856

de Guerrero, M. (2018). Going covert: Inner and private speech in language learning. *Language Teaching*, 51, 1–35. https://doi.org/10.1017/S0261444817000295

De Palma, M.-J., & Ringer, J. M. (2011). Toward a theory of adaptive transfer: Expanding disciplinary discussions of ''transfer'' in second-language writing and composition studies. *Journal of Second Language Writing*, 20, 134–147. https://doi.org/10.1016/j.jslw.2011.02.003

Dong, Y., Gui, S., & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition*, 8, 221–238. https://doi.org/10.1017/S1366728905002270

Dressman, M., Wilder, P. & Connor, J. (2005). Theories of failure and the failure of theories: A cognitive/sociocultural/macrostructural study of eight struggling students. *Research in the Teaching of English*, 40, 8–61.

Duff, P. (2010). Language socialization into academic discourse communities. *Annual Review of Applied Linguistics*, 30, 169–192. https://doi.org/10.1017/S0267190510000048

Ellis, N. (2019). Essentials of a theory of language cognition. *Modern Language Journal*, 103, Supplement, 39–60. https://doi.org/10.1111/modl.12532

Engeström, Y. (2008). *From teams to knots: Activity-theoretical studies of collaboration and learning at work*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511619847

Engeström, Y. (2015). *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). Cambridge: CUP.

Ferris, D. (2012). Written corrective feedback in second language acquisition and writing studies. *Language Teaching*, 45, 446–459. https://doi.org/10.1017/S0261444812000250

Fitzgerald, J. (2006). Multilingual writing in preschool through 12th grade. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 337–354). New York, NY: Guilford Press.

Flowerdew, J., & Li, Y. (2007). Language re-use among Chinese apprentice scientists writing for publication. *Applied Linguistics*, 28, 440–465. https://doi.org/10.1093/applin/amm031

Fogal, G. (2017). Tracking microgenetic changes in authorial voice development from a complexity theory perspective. *Applied Linguistics*, advance view. https://doi.org/10.1093/applin/amx031

Gentil, G. (2005). Commitments to academic biliteracy: Case studies of francophone university writers. *Written Communication*, 22, 421–471. https://doi.org/10.1177/0741088305280350

Gere, A., Limlamai, N., Wilson, E., Saylor, K., & Pugh, R. (2018). Writing and conceptual learning in science: An analysis of assignments. *Written Communication*, 36, 99–135. https://doi.org/10.1177/0741088318804820

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81, 710–744. https://doi.org/10.17763/haer.81.4.t2kom13756113566

Hall, C. (1990). Managing the complexity of revising across languages. *TESOL Quarterly*, 24, 43–60. https://doi.org/10.2307/3586851

Haneda, M. (2007). Modes of engagement in second language writing: An Activity Theoretical perspective. *Canadian Modern Language Review*, 64, 301–331. https://doi.org/10.3138/cmlr.64.2.297

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11, 329–350. https://doi.org/10.1016/S1060-3743(02)00091-7

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29, 369–388. https://doi.org/10.1177/0741088312451260

van Heuven, W., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Research Reviews*, 64, 104–122. https://doi.org/10.1016/j.brainresrev.2010.03.002

Hirvela, A. (2011). Writing to learn in content areas: Research insights. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 38–59). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.06hir

Hornberger, N. (1989). Continua of biliteracy. *Review of Educational Research*, 59, 271–296. https://doi.org/10.3102/00346543059003271

Hornberger, N. (Ed.). (2003). *Continua of biliteracy: An ecological framework for educational policy, research, and practice in multilingual settings*. Clevedon, UK: Multilingual Matters. https://doi.org/10.21832/9781853596568

Ivanic, R. (1998). *Writing and identity: The discoursal construction of identity in academic writing*. Amsterdam: John Benjamins. https://doi.org/10.1075/swll.5

James, M. A. (2010). An investigation of learning transfer in English-for-general-academic-purposes writing instruction. *Journal of Second Language Writing*, 19(4), 183–206. https://doi.org/10.1016/j.jslw.2010.09.003

James, M. A. (2014). Learning transfer in English-for-academic-purposes contexts: A systematic review of research. *Journal of English for Academic Purposes*, 14, 1–13. https://doi.org/10.1016/j.jeap.2013.10.007

Janks, H. (2010). *Literacy and power*. New York, NY: Routledge.

Jun, S. W., Ramírez, G., & Cumming, A. (2010). Tutoring adolescents in literacy: A meta-analysis. *McGill Journal of Education*, 45(2), 219–238. https://doi.org/10.7202/045605ar

Kang, J. Y. (2005). Written narratives as an index of L2 competence in Korean EFL learners. *Journal of Second Language Writing*, 14, 259–279. https://doi.org/10.1016/j.jslw.2005.10.002

Kang, E.-Y., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, 99, 1–18. https://doi.org/10.1111/modl.12189

Kaplan, R. (1966). Cultural thought patterns in intercultural education. *Language Learning*, 16, 1–20. https://doi.org/10.1111/j.1467-1770.1966.tb00804.x

Kibler, A. (2014). From high school to the noviciado: An adolescent linguistic minority student's multilingual journey in writing. *Modern Language Journal*, 98, 629–651. https://doi.org/10.1111/modl.12090

Klein, P., Arcon, N., & Baker, S. (2016). Writing to learn. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 243–256). New York, NY: Guilford Press.

Knutson, E. (2006). Thinking in English, writing in French. *The French Review*, 80, 88–109.

Kobayashi, H., & Rinnert, C. (2012). Understanding L2 writing development from a multicompetence perspective. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 101–134). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781934078303.101

Kroll, J., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25, 497–514. https://doi.org/10.1080/20445911.2013.799170

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: OUP.

Larsen-Freeman, D. (2013). Transfer of learning transformed. *Language Learning*, 63, Supplement 1, 107–129. https://doi.org/10.1111/j.1467-9922.2012.00740.x

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: OUP.

Lee, I. (2013). Research into practice: Written corrective feedback. *Language Teaching*, 46, 108–119. https://doi.org/10.1017/S0261444812000390

Lei, X. (2008). Exploring a sociocultural approach to writing strategy research: Mediated actions in writing activities. *Journal of Second Language Writing*, 17, 217–236. https://doi.org/10.1016/j.jslw.2008.04.001

Leki, I. (2007). *Undergraduates in a second language: Challenges and complexities of academic literacy development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Leki, I., Cumming, A., & Silva, T. (2008). *A synthesis of research on second language writing in English*. London: Routledge.

Li, M., & Zhu, W. (2017). Good or bad collaborative wiki writing: Exploring links between group interactions and writing products. *Journal of Second Language Writing*, 35, 38–53. https://doi.org/10.1016/j.jslw.2017.01.003

Li, Y., & Casanave, C. (2012). Two first-year students' strategies for writing from sources: Patchwriting or plagiarism? *Journal of Second Language Writing*, 21, 165–180. https://doi.org/10.1016/j.jslw.2012.03.002

Linck, J., Hoshino, N., & Kroll, J. (2008). Cross-language lexical processes and inhibitory control. *Mental Lexicon*, 3, 349–374. https://doi.org/10.1075/ml.3.3.06lin

MacArthur, C., & Graham, S. (2016). Writing research from a cognitive perspective. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 24–40). New York, NY: Guilford Press.

Macqueen, S. (2012). *The emergence of patterns in second language writing*. Bern, Switzerland: Peter Lang. https://doi.org/10.3726/978-3-0351-0437-0

Manchón, R. M. (Ed.). (2011). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31

McCarthy Young, K., & Leinhardt, G. (1998). Writing from primary documents: A way of knowing in history. *Written Communication*, 15, 25–68. https://doi.org/10.1177/0741088398015001002

Mizumoto, A., Hamatani, S., & Imao, Y. (2017). Applying the bundle–move connection approach to the development of an online writing support tool for research articles. *Language Learning*, 67, 885–921. https://doi.org/10.1111/lang.12250

Murphy, L., & Roca de Larios, J. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL learners. *Journal of Second Language Writing*, 19, 61–81. https://doi.org/10.1016/j.jslw.2010.02.001

Newell, G., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46, 273–304.

Nicolás-Conesa, F., Roca de Larios, J., & Coyle, Y. (2014). Development of EFL students' mental models of writing and their effects on performance. *Journal of Second Language Writing*, 24, 1–19. https://doi.org/10.1016/j.jslw.2014.02.004

Norton, B. (2013). *Identity and language learning: Extending the conversation*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781783090563

Norton, B., & Toohey, K. (2011). Identity, language learning, and social change. *Language Teaching*, 44, 412–446. https://doi.org/10.1017/S0261444811000309

Park, J-H., & De Costa, P. (2015). Reframing graduate student writing strategies from an Activity Theory perspective. *Language and Sociocultural Theory*, 2, 25–50. https://doi.org/10.1558/lst.v2i1.24977

Parks, S., & Maguire, M. (1999). Coping with on-the-job writing in ESL: A constructivist-semiotic perspective. *Language Learning*, 49, 143–175. https://doi.org/10.1111/1467-9922.00073

Pérez, B. (2004). *Becoming biliterate: A study of two-way bilingual immersion education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Prior, P. (2006). A sociocultural theory of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54–66). New York, NY: Guilford Press.

Qi, D. (1998). An inquiry into language-switching in second language composing processes. *Canadian Modern Language Review*, 54, 413–435. https://doi.org/10.3138/cmlr.54.3.413

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon, UK: Multilingual Matters.

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Bristol, UK: Multilingual Matters.

Robinson, P. (Ed.). (2011). *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance*. Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.2

Roca de Larios, J., Murphy, L., & Manchón, R. M. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a foreign language. *Journal of Second Language Writing*, 8, 13–44. https://doi.org/10.1016/S1060-3743(99)80111-8

Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' writing ability and motivation. *TESOL Quarterly*, 45, 81–105. https://doi.org/10.5054/tq.2011.240861

Schachter, J., & Celce-Murica, M. (1977). Some reservations concerning error analysis. *TESOL Quarterly*, 11, 441–451. https://doi.org/10.2307/3585740

Schleppegrell, M. (1994). *The language of schooling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schleppegrell, M. (2016). Content-based language teaching with functional grammar in the elementary school. *Language Teaching*, 49, 116–128. https://doi.org/10.1017/S0261444814000093

Shi, L. (2010). Textual appropriation and citing behaviors of university undergraduates. *Applied Linguistics*, 31, 1–24. https://doi.org/10.1093/applin/amn045

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing. https://doi.org/10.1037/11256-000

Stein, P. (2008). *Multimodal pedagogies in diverse classrooms: Representation, rights and resources*. London: Routledge.

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847699954

Strijkers, K. (2016). A neural assembly-based view on word production: The bilingual test case. *Language Learning*, 66, Supplement 2, 92–131. https://doi.org/10.1111/lang.12191

Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Brynes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London: Continuum.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step toward second language learning. *Applied Linguistics*, 16, 371–391. https://doi.org/10.1093/applin/16.3.371

Tardy, C. (2009). *Building genre knowledge*. Anderson, SC: Parlour Press.

Tardy, C. (2016). *Beyond convention: Genre innovation in academic writing*. Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.5173647

To, J., & Carless, D. (2015). Making productive use of exemplars: Peer discussion and student guidance for positive transfer of strategies. *Journal of Further and Higher Education*, 40, 746–764. https://doi.org/10.1080/0309877X.2015.1014317

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Uzawa, K., & Cumming, A. (1989). Writing strategies in Japanese as a foreign language: Lowering or keeping up the standards. *Canadian Modern Language Review*, 46, 178–194. https://doi.org/10.3138/cmlr.46.1.178

Verhoeven, L. (1994). Transfer in bilingual development: The linguistic development hypothesis revisited. *Language Learning*, 44, 381–415. https://doi.org/10.1111/j.1467-1770.1994.tb01112.x

Verspoor, M., Schmid, M., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21, 239–263. https://doi.org/10.1016/j.jslw.2012.03.007

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wang, L. (2003). Switching to first language among writers with differing second-language proficiency. *Journal of Second Language Writing* 12, 347–375. https://doi.org/10.1016/j.jslw.2003.08.003

Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL learners. *Journal of Second Language Writing*, 11, 225–246. https://doi.org/10.1016/S1060-3743(02)00084-X

Wardaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly*, 4, 123–130. https://doi.org/10.2307/3586182

Wertsch, J. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.

Whalen, K., & Menard, N. (1995). L1 and L2 writers' strategic and linguistic knowledge: A model of multiple-level discourse processing. *Language Learning*, 45, 381–418. https://doi.org/10.1111/j.1467-1770.1995.tb00447.x

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–333. https://doi.org/10.1016/j.jslw.2012.09.007

Wilson, J. S. (2013). The role of social relationships in the writing of multilingual adolescents. In L. De Olieviera & T. Silva (Eds.), *L2 writing in secondary classrooms* (pp. 87–103). New York, NY: Routledge.

Winsor, D. (2003). *Writing power: Communication in an engineering center*. Albany, NY: State University of New York Press.

A dual-process model of L1 writing processes

Implications for L2 writing research agendas on processing and language development

David Galbraith and Zulaikha Al-Saadi
University of Southampton

Current models of the cognitive processes involved in L1 writing treat them as problem-solving processes, and characterize expert / novice differences as a contrast between a knowledge-transforming approach and a knowledge-telling approach. Empirical research informed by these models has typically used this distinction as a characterization of the processes involved but has not explicitly tested what effect these have on the writer's knowledge. In this chapter, we argue that these models neglect the implicit nature of much of the writer's knowledge, and present an alternative dual-process model of writing, which is designed to take this into account. We then describe recent studies in L1 that support the main claims of the dual-process model, call into question some of the assumptions of problem-solving models and suggest that an alternative rough-drafting strategy may help overcome the conflict between text quality and the development of the writer's understanding. We conclude by discussing the implications for research into L2 writing processes and drafting strategies.

Classical cognitive models of the writing process in L1 typically assume that writing is a matter of translating preconceived ideas into text (Hayes & Flower, 1980a). Although these models recognize that the writer develops their understanding during writing, this is assumed to be a consequence of adapting pre-existing ideas to the specific rhetorical context for writing. In such models, text production is treated primarily as an impediment to the operation of the higher-level thinking processes involved in knowledge transforming, and strategies such as outlining, which separate idea generation from the process of text production, are recommended for improving the quality of text (Kellogg, 1990). In this paper, we will argue that these models have neglected the implicit nature of knowledge-representation and, hence, the role that text production plays in enabling the writer to constitute their implicit knowledge in the text. This links with the dimension of "learning by writing" outlined in the overall framework presented by Cumming (Chapter 2). We

will describe an alternative, dual-process model of writing (Galbraith & Baaijen, 2018), in which writing is the joint product of two conflicting processes: An explicit problem-solving process, and an implicitly-controlled knowledge constituting process, taking place during the formulation of thought in language, and responsible for developing the writer's understanding of a topic. We will then present the findings from recent empirical research that provide evidence for these two conflicting components of the writing process. We will conclude by discussing the implications for L2 writing.

Writing as problem solving

In the classic model of cognitive processes in writing (Hayes & Flower, 1980), writing was treated as a problem-solving activity involving the three fundamentally different processes of planning, translating, and reviewing. Planning was treated as a conceptual process, which involved goal setting and the generation and organization of ideas to satisfy those goals; translating was treated as a linguistic process, which involved the translation of those ideas into words; and reviewing was treated as a process which combined linguistic and conceptual operations, and involved reading and editing previously generated text. These three processes operated on information stored in long-term memory about the topic, audience and writing plans, and in the context of the task environment, which consisted of the writing assignment (the topic, audience, and motivating cues) and the text produced so far. A key feature of the model was that, although the fundamental processes were labelled in the same way as a traditional stage model of writing (plan-write-revise), they were not assumed to correspond to stages in the writing process, but rather to refer to different cognitive operations that could occur at any time during writing. Individual differences in how writers combined these operations were represented by a monitor, which varied in how it was configured.

Research informed by this model – often using verbal protocols to identify processes and comparing expert and novice protocols – led to two broad conclusions (Hayes, 1996; Hayes & Flower, 1986). First, a fundamental difference between experts and novices arises from the goals towards which writing is directed. Novices' fundamental goal is to express their knowledge about a topic in words. To use Bereiter and Scardamalia's (1987) terms, they take a *knowledge-telling* approach to writing, in which ideas are retrieved from long-term memory and translated directly into words. This involves relatively little planning, and reviewing is restricted to evaluating surface features of the text. The result is what Flower (1979) called "writer-based" prose. By contrast, more expert writers' fundamental goal is to design a text that communicates effectively with their audience. Accordingly, they

take a *knowledge-transforming* approach to writing, in which both the retrieval and evaluation of ideas and the revision of text are guided by rhetorical goals. This involves more elaborate planning, both before and during writing, and more extensive revision of content, rather than simply of surface features of the text. The result is what Flower (1979) called "reader-based" prose. A key feature of this approach, as its name implies, is that, in trying to produce more communicatively-effective text, the writer is also forced to re-evaluate their knowledge.

The second assumption of the problem-solving model was that a fundamental conflict in writing lay in cognitive overload arising from the need to combine writing processes within a limited-capacity system (Flower & Hayes, 1980b). This prompted a range of work investigating the relationship between working memory capacity and writing performance (Kellogg, 1996; Kellogg, Whiteford, Turner, Cahill, & Mertens, 2013; McCutchen, 2000) and led to the explicit incorporation of working memory within Hayes and Flower's model (Hayes, 1996). An important consequence of this was a more precise characterization of drafting strategies, not simply as a description of alternative possible ways of combining the basic writing processes, but rather as having the function of managing cognitive overload. In a series of experiments, Kellogg (1988, 1990) tested this overload hypothesis by comparing the relative effectiveness for text quality of *outline planning* – designed to reduce cognitive overload by separating the generation and organization of ideas from full text production – and *rough drafting* – designed to reduce cognitive overload by separating translation from reviewing – with writing a *single draft.* Kellogg (1990) concluded that the findings provided clear support for the overload hypothesis and convincing experimental evidence for the benefits of outlining.

In combination, these features of the problem-solving model suggest that effective writing in L1 depends on the extent to which writing is directed towards rhetorical goals and on outline planning to reduce cognitive overload. When applied effectively, the result should be more effective text and the development of the writer's understanding of the topic. Reviews of subsequent research (Graham, 2006; Graham & Perin, 2007) suggest that writing instruction informed by these principles consistently helps students write more effectively.

There have been important developments to Hayes and Flower's (1980) model since its inception. In Hayes's (2012) most recent update of the model, the basic processes have become less monolithic, and map less directly onto traditional stage models of writing. This reconceptualization of the model provides a clearer distinction between general cognitive functions (control processes and general cognitive resources) and more specific written-language production processes (which include a distinction between the linguistic formulation of ideas and transcription of language into written form). However, despite these changes, the central importance of the writer's goals and the writer's ability to manage cognitive overload is retained.

In particular, as we discuss further below, the writer's long-term memory is still treated as an undifferentiated source of content.

Research into the cognitive processes in L2 writing (see Roca de Larios, Nicolás-Conesa, & Coyle, 2018, for a review) has tended to make the same assumptions about the basic processes involved. Reviews of such cognitively-inspired research (e.g. Cumming, 2016; Fitzgerald, 2006; Roca de Larios et al. 2018) have typically concluded that similar kinds of problem-solving processes occur in L1 and L2 writing, but that these are constrained by L2 proficiency (Stevenson, Schoonen, & de Glopper, 2006).

Dual process model of writing in L1

One of the striking features of cognitive models of writing is how little they have to say about how the writer's knowledge is represented. This has continued to be characterized simply as an undifferentiated store of information in long-term memory (Hayes, 2012; Hayes & Flower, 1980). Knowledge in long-term memory is assumed to be explicitly represented as declarative knowledge and to be retrieved from memory before being translated into words. An important consequence is that research has focused on the processes involved in complying with external communicative constraints rather than on those involved in capturing the writer's own distinctive view of the topic.

The dual-process model (Galbraith, 2009; Galbraith & Baaijen, 2018), by contrast, emphasizes the implicit nature of much of our knowledge and distinguishes between explicit and implicit memory systems. This is based on the well-established Complementary Learning Systems (CLS) theory (Kumaran, Hassabis, & McClelland, 2016; McClelland, McNaughton, & O'Reilly, 1995; O'Reilly, Bhattacharyya, Howard, & Ketz, 2014). CLS theory postulates two memory systems: A semantic system, located in the neocortex and characterized by O'Reilly et al. (2014) "as a distributed, overlapping system for gradually integrating across episodes to extract latent semantic structure" (p. 1229), and an episodic system, located in the hippocampus and characterized "as a sparse, pattern-separated system for rapidly learning episodic memories" (p. 1229). The dual-process model claims that problem solving models capture the processes involved in manipulating explicit information retrieved from the episodic memory system, but neglect the processes involved in synthesizing content according to the structure of the semantic memory system. It therefore claims that two distinctive processes are involved in generating content during writing, and that these make different contributions to the development of the writer's understanding of the topic and the quality of the text that they produce. The application of these tenets to L2 writing will be discussed in a later section.

The knowledge-transforming process

According to the dual-process model (Galbraith, 2009; Galbraith & Baaijen, 2018), the knowledge-transforming process operates on individual ideas retrieved from episodic memory, which may be evaluated and manipulated in working memory to satisfy the writer's goals. It operates most efficiently when ideas are represented in a fixed, abbreviated form so that the limited capacity working memory system can focus its resources on evaluating their contribution to the writer's goals for the text. When the writer's goals are simply to translate their ideas into words, as in novice writing, the result is what Bereiter and Scardamalia (1987) describe as knowledge telling. When the writer has more elaborate goals for communicating with their readers, as in more expert writing, search of episodic memory is more strategic, and evaluation of potential content based on the writer's goals results in knowledge transformation. However, according to the dual-process model, this does not, by itself, involve the generation of novel content, but rather is restricted to the reorganization of existing content. To the extent that this leads to the formation of a more coherent object in episodic memory, this will contribute to the development of the writer's knowledge and, since it involves the adaptation of content to communicative goals, will be associated with better quality text.

The knowledge-constituting process

The characteristics of the knowledge-constituting process were first formulated in Galbraith (1999). In its current form (Galbraith & Baaijen, 2018), the process is assumed to have three crucial features. First, knowledge within the semantic memory system is represented implicitly by the strength of the connections between units within a constraint satisfaction network. The strengths of these fixed connections are the product of an individual's learning history and reflect the totality of an individual's experience. Galbraith (1999) refers to this intrinsic organization of the writer's knowledge as the writer's *disposition* towards the topic. The content of this implicit disposition is synthesized in response to a topic by constraint satisfaction within the network, and is revealed as an output to be expressed in language. Second, this initial synthesis of content in response to the topic is only a partial "best fit" to the writer's disposition, constrained by the limited capacity of the language system. It is not a direct translation of pre-existing content. Third, inhibitory feedback from this initial output provides a new input to the disposition, which prompts a further cycle of constraint satisfaction within the network. This has the effect of reducing the activation of units associated with the initial message, and prompts the synthesis of previously suppressed content corresponding to the "remainder" of the content. In combination, these three features enable the writer's implicit

understanding of the topic to be discursively constituted in the text. Crucially, this is not a matter of retrieving a series of pre-stored propositions. Each successive proposition depends on the output of the preceding synthesis and in turn influences the synthesis of its successor.

According to the dual process model, then, the knowledge-constituting process is intrinsically a process of discovery. It operates best when writers synthesize their thought in explicit, connected propositions and when successive propositions are produced as dispositional responses to preceding text. Crucially, it depends on the writer's thought being allowed to unfold, guided by the implicit organization of their disposition towards the topic, without interruption by external goals. It will lead to the development of the writer's understanding when the content that it produces does not correspond to existing content stored in episodic memory. However, because this depends on the extent to which content is dispositionally produced, rather than in response to rhetorical goals, there will no necessary relationship between the development of understanding and the quality of the text.

The two processes in combination

In principle, these two processes are assumed to have complementary functions. The knowledge-transforming (problem-solving) process is designed to ensure that content is coherently organized to satisfy rhetorical goals, and is assumed to guide the search and manipulation of existing content retrieved from episodic memory. When existing content is not available in episodic memory, the knowledge-transforming process can set this as a goal for the knowledge-constituting process. The knowledge-constituting process is designed to synthesize content reflecting the writer's implicit understanding of the topic, and produces output, including novel content, which is stored in episodic memory, to be organized by the knowledge-transforming process. In combination, the two processes lead to the creation of a coherent knowledge object, which reflects the writer's implicit understanding of the topic and satisfies their rhetorical goals.

There are two key differences between problem-solving models of writing and the dual-process model. First, the problem-solving model assumes that the development of understanding depends on the same deliberate, rhetorically-guided problem-solving processes applied to both text production and the higher-level processes involved in global planning and revision. Hence the fundamental contrast in writing is between a knowledge-telling approach, associated with less development of the writer's understanding and the production of less communicatively effective text, and a knowledge-transforming approach, associated with both greater development of the writer's understanding and the production of more communicatively-effective text. By contrast, the dual-process model assumes that

the development of the writer's understanding depends on two different processes with different relationships with text quality. The knowledge-constituting process leads to the constitution of the writer's implicit understanding in explicit text, but because it is dispositionally-guided rather than rhetorically organized, will be associated with less communicatively-effective text. The knowledge-transforming process leads to the reorganization of the writer's knowledge to satisfy rhetorical goals, and so develops both the writer's knowledge of the global structure of the text and contributes to the communicative effectiveness of the text.

The second difference is over the nature of the fundamental conflict in writing. For problem solving models, this is cognitive overload arising from the range of different processes that have to be combined during writing. In principle, this can be reduced by a range of different drafting strategies; in practice, the evidence to date suggests that the most effective drafting strategy is outlining (Kellogg, 1988, 1990), which enables the writer to focus on content generation and organization first, before devoting attention to the translation of the pre-determined content into words. By contrast, the dual-process model assumes that the fundamental conflict is between the two different sources of organization involved in the knowledge-constituting and knowledge-transforming processes: The implicit organization of the writer's disposition, which guides the emergence of text during the knowledge-constituting process, and explicit organization of the text to satisfy rhetorical goals. Accordingly, it assumes that although outline planning may have benefits for text quality, it will do so at the expense of capturing the writer's distinctive understanding of the topic. It suggests, instead, that a revision strategy, in which an initial, dispositionally-guided draft of text is revised over a series of drafts into a rhetorically appropriate form, will enable the writer to both capture their implicit understanding of the topic and to present it in a rhetorically effective form.

Evidence for the dual process model

Research on L1and L2 writing inspired by problem-solving models has typically relied on verbal protocols to provide information about writing processes, with the result that the focus of research has been on variations in the higher-level thinking processes involved in writing, rather than on variations in how text production has been carried out. The exception to this has been a series of studies by Hayes and his colleagues (Chenoweth & Hayes, 2003; Hayes, 2009) which, though often reliant on verbal protocols, have investigated the processes involved in text production in more detail. The most important finding from this research has been that writers produce language in shorter bursts in their L2 than in L1. More recently, however, keystroke logging has been used to provide more detail about

the moment-by-moment processes involved in producing text (see Lindgren & Sullivan, 2019; Révész & Michel, 2019, for recent collections of studies on both L1 and L2 writing using keystroke logging). Somewhat surprisingly, there has been very little research designed to test the problem-solving models' assumptions about the relationship between writing processes and the development of the writer's understanding.[1] It has been taken for granted that the more elaborate processes characteristic of the knowledge-transforming process do in fact lead to a development of the writer's understanding. In this section, we describe the results of a recent study explicitly designed to test the competing claims of the problem-solving and dual-process models (see Galbraith & Baaijen, 2018, for a review of the empirical basis for the dual-process model more generally). We then consider some studies examining the relative effectiveness of different drafting strategies.

Keystroke studies of text production

Baaijen and Galbraith (2018) used keystroke-logging to identify the relationship of two components of the writing process – sentence production and global linearity – with text quality and the development of understanding through writing. Seventy eight undergraduates from a Dutch university were asked to write an article suitable for the university newspaper (in their L1) discussing whether "our growing dependence on computers and the Internet is a good development or not". They were then randomly assigned to one of two conditions: (i) an *outline* planning condition, in which they were given 10 minutes to generate a list of ideas, followed by 5 minutes to create an organized outline of the text to be written; or (ii) a *synthetic* planning condition, in which, following 10 minutes for idea generation, they were given 5 minutes to sum up their overall opinion of the topic in a single sentence. This was designed to manipulate the extent to which the writers had to create an explicit organization for their text before writing. Both groups were then given 30 minutes to write a well-structured article for the university newspaper. Keystrokes during this writing period were recorded using Inputlog (Leijten & van Waes, 2013). To measure the effect of writing on their understanding of the topic, both groups were asked to rate how much they felt they knew about the topic on a 7-point scale before and after writing. Text quality was assessed by two independent raters on a 9-point holistic scale. The raters were asked make an overall judgement based on the coherence of the argument, and the originality and appropriateness of tone of the article (Baaijen & Galbraith, 2018, p. 209).

The two writing process measures were constructed as orthogonally related, composite measures derived from individual features of the keystroke logs, and

---

1.   There has been extensive research on writing to learn content but this has not specifically investigated effects on the writer's subjective understanding of a topic.

were designed to capture the two components of the dual-process model (Baaijen, Galbraith, & de Glopper, 2012). These represented independent dimensions reflecting how individual sentences were produced and how they were sequenced. The *sentence production* dimension distinguished between, at one extreme, controlled sentence production – relatively lengthy pauses between sentences followed by clean bursts of text, with little revision – and, at the other extreme, spontaneous sentence production – brief pauses between sentences followed by extensively revised bursts of text. The *global linearity* dimension distinguished between linear text production at one extreme – sentences were produced one after the other with little evidence of recursion – and non-linear text production at the other extreme – sentences were produced recursively, with greater evidence of insertions within, and revisions of, previously produced text.

   This study had three important findings. First, consistent with the dual-process model, the two process measures made independent contributions to increases in understanding after writing. Global linearity was negatively related to increased understanding. This is consistent with the distinction between knowledge-telling and knowledge-transforming made by both the dual-process model and problem-solving models: Linearly produced texts were associated with relatively little increased understanding; texts including more extensive revision of global structure were associated with greater increases in understanding. Of particular interest, however, was the relationship with the sentence production measure, which is shown in Figure 1 (sentence production is plotted 1SD above and below the mean score on this dimension).
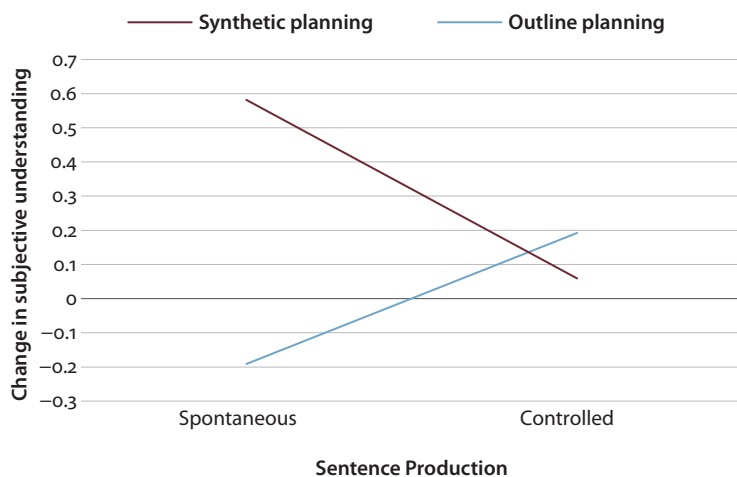

Figure 1. Relationship between sentence production and change in understanding as a function of type of planning (Baaijen & Galbraith, 2018). Reproduced with permission

As can be seen in Figure 1, the relationship depended on the type of planning carried out in advance of writing. When writing was synthetically planned (i.e. no explicit organization was imposed on the text before writing), increases in understanding were associated with more spontaneously produced sentences, and declined towards zero the more controlled sentence production was. By contrast, when writing was outline planned, increases in understanding were close to zero, and were at a minimum for texts consisting of more spontaneously produced sentences. These results closely correspond with the dual-process model's predictions: Increased understanding is associated with spontaneous rather than controlled sentence production, and is suppressed when writing is outline planned. Baaijen and Galbraith (2018) speculated that the marked difference between synthetic and outline planning reflected the dual-process model's claim that content is dispositionally synthesized in the synthetic planning condition but retrieved directly from episodic memory in the outline planning condition. Although this claim doesn't follow directly from the keystroke measures, which don't provide information about the underlying cognitive processes, it could be tested in future research by examining the extent to which neocortical and hippocampal networks are activated under different planning conditions.

Overall, the relationships between the keystroke indicators of writing processes and changes in the writers' subjective understanding are compatible with the dual process model's claims: Two different kinds of process are associated with the development of the writer's understanding, and characteristics of the relationship with sentence production are consistent with the claims made for the knowledge-constituting process. The next question is how these processes are related to text quality. Baaijen and Galbraith (2018) found that these varied depending on the type of planning carried out before writing. In the outline planning condition, the writing process measures were unrelated to the quality of the text, which Baaijen and Galbraith suggested was because the quality of the text depended primarily on the quality of the planning carried out during outlining. By contrast, in the synthetic planning condition, there were clear relationships between the two writing process measures and text quality: Higher quality text was produced by writers who combined more global revision during writing with more controlled sentence production. Furthermore, the poorest quality text was produced precisely by the writers who wrote most spontaneously in this condition. These results, therefore, strongly confirmed the dual-process model's claim that the two processes have differential relationships with text quality. Synthetically planned, spontaneous sentence production is distinctively associated with the development of the writer's understanding but is, as the dual-process model predicts, negatively related to text quality. By contrast, global revision is both positively associated with

the development of understanding, and also, in the synthetic planning condition at least, positively associated with text quality.

The final important finding of this study emerged from Baaijen and Galbraith's analysis of the relationships between text quality and the development of the writer's understanding in the different planning conditions. Adding text quality and its associated interactions as predictors to the model of the development of understanding increased the proportion of the variance accounted for substantially, suggesting that the relationship between text quality and the development of understanding has an important added role in the process. Although the causal direction of this relationship is difficult to disentangle, two features were clear. First, the relationship between synthetically planned, spontaneous sentence production and increased understanding was even stronger when text quality and its associated interactions were controlled for. This confirms the contrasting relationships that the knowledge-constituting process has with the development of understanding and text quality. Second, the relationship of global revision with the development of understanding and text quality was more complicated than the initial analysis had suggested. These relationships are shown in Figure 2. (Text quality is plotted at 1SD above and below the mean score for this variable, and high (+1SD) and low (−1SD) levels of global linearity are plotted as separate lines on the graph).

As can be seen in Figure 2, for high levels of global revision (non-linear GL), the relationship between text quality and development of understanding is in opposite directions in the two planning conditions. In the outline planning condition, the relationship is negative: High text quality is associated with zero change in understanding; and when changes in understanding do occur, they are associated with poor text quality. This calls strongly into question the general assumption that expert, high quality writing is a knowledge-transforming process. In this condition at least, high quality writing appears to depend precisely on writing <u>not</u> leading to the transformation of knowledge. By contrast, in the synthetic planning condition, where developments in understanding are generally higher than in outline planning, global revision is associated with increases in understanding to the extent that it is also associated with high quality text.

Baaijen and Galbraith (2018) suggest that one possible explanation for these contrasting relationships may lie in the different goals towards which revision is directed in the two planning conditions. In outline-planned writing, revision may be designed to revise the text to ensure that it complies with the writer's initial plan. When this is carried out successfully, text quality improves but does not lead to the development of the writer's understanding; when it is not carried out successfully, the writer has to modify their original plan – globally revising their text – which leads to a development in their understanding, but at the expense of maintaining the quality

a.

b.
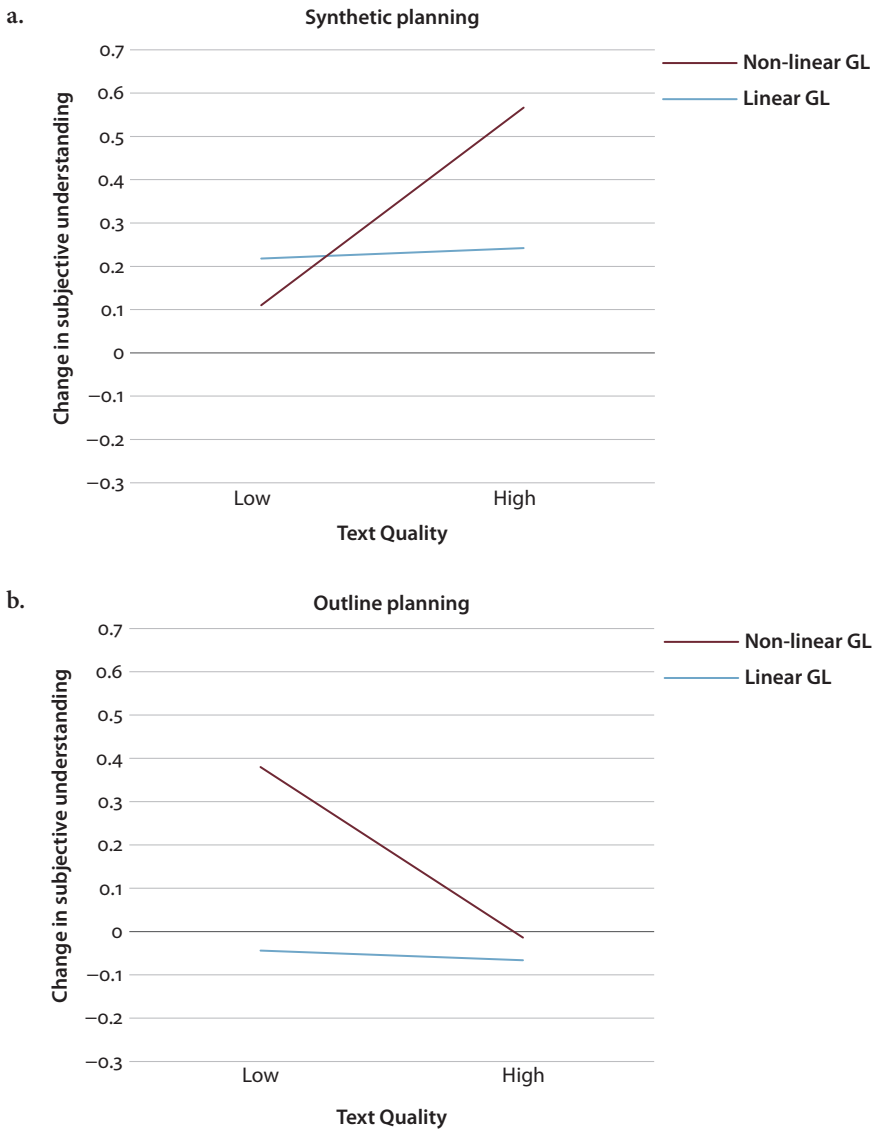
Figure 2. Relationship between text quality and change in subjective understanding as a function of global linearity and type of planning (Baaijen & Galbraith, 2018, p. 214). Reproduced with permission

of the text. By contrast, in the synthetic planning condition, where the development of understanding through spontaneous sentence production is at a maximum, and there is no pre-determined structure, global revision may be designed to construct a global structure compatible with the content as it emerges in the text. When this is

successful, text quality is improved and the writer develops a better understanding of the global structure of their thought; when it is unsuccessful, the writer's understanding is not enhanced and the quality of the text is poor. This, in effect, would preserve the problem-solving account of the role of the knowledge-transforming process in developing knowledge and text quality, but would contextualise it in terms of the writer's planning strategy and their goals in revision.

Drafting strategies

The writing studied by Baaijen and Galbraith (2018) was carried out in a single, half-hour draft, and writers were required to produce a well-structured article in that time, with relatively little time allowed for constructing an outline. A natural question is whether allowing writers to write multiple drafts, with greater time for revision, might enable them to produce better quality text, and alter the relationship between the processes carried out in the initial draft and the quality of the final text. Previous research by Kellogg (1988, 1990, 1994) has suggested that outlining is the most effective form of drafting strategy, and that rough drafting strategies, in which text production was separate from revision, were relatively ineffective. However, Galbraith and Torrance (2004) suggested that this failure to find an effect lay in the form of rough drafting strategy that had been implemented in these studies, and that the key ingredient of an effective rough drafting strategy was that content generation during text production should be separated from the need to organize the text in terms of rhetorical goals. In terms of the dual-process model, this amounts to suggesting that the initial draft should be dispositionally controlled and text production should follow the path of thought as it unfolds.

In the experiment designed to test this, writing was divided into three phases. In the first phase, which lasted 20 minutes, they manipulated how organized the text had to be and whether it had to be expressed in connected sentences or in note-form. This resulted in four initial drafting conditions: (i) organized notes; (ii) organized sentences; (iii) unorganized notes; and (iv) unorganized sentences. Following this phase, all the writers were given 5 minutes to read over the initial draft, noting down any major changes they wanted to make, and then writing down a single sentence summing up their main point. In the final, 20-minute phase, the writers were asked to produce a final, well-formed draft of the text under one of two conditions: either (i) with initial draft available for reference or (ii) with the initial draft removed, so that they had to produce a fresh draft of text.

The key finding of this study was that the quality of the final text varied depending on how organized the initial draft was required to be and on whether or not it was removed before the final draft was written. Figure 3a shows the ratings of quality for the final drafts produced in each condition.

a.

Galbraith & Torrance (2004)

b.
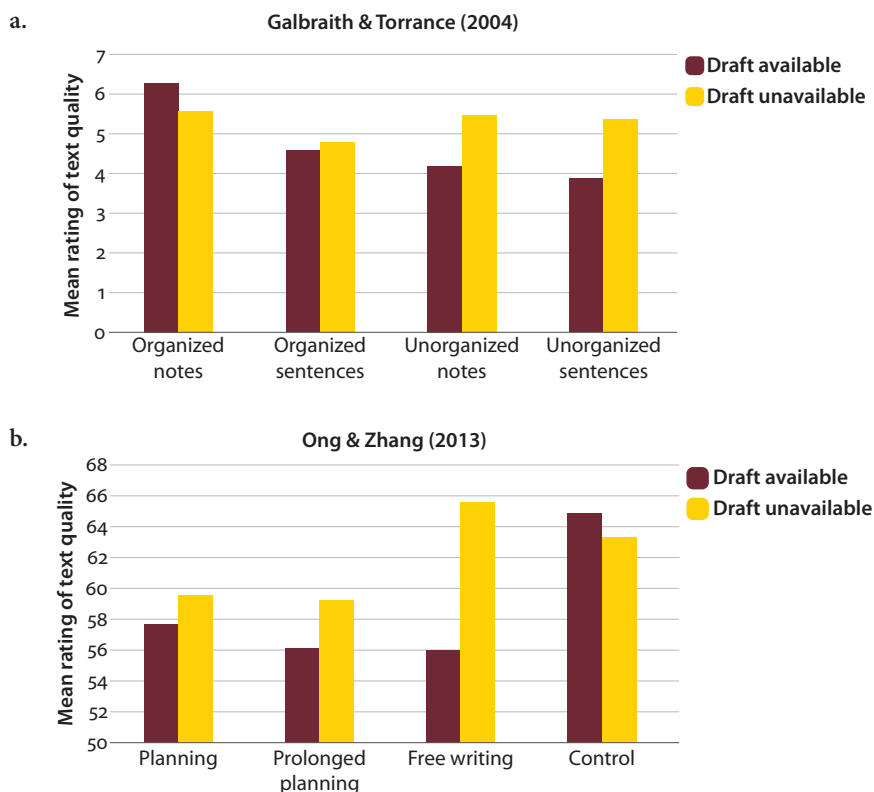
Ong & Zhang (2013)

Figure 3. Relationship between drafting condition and text quality in L1 (Galbraith & Torrance, 2004) and L2 (Ong & Zhang, 2013) (adapted from data presented in original papers)

As Figure 3a shows, when the initial draft remained available for reference during the final draft (the purple bars), the results essentially replicated Kellogg's previous research. When writers were able to generate and organize their ideas in note-form before writing the final text, they were able to produce better texts than when the initial draft involved producing full text at the same time as generating and organizing ideas, or when writers were asked to produce unorganized initial drafts in either note-form or full sentences. However, when the initial draft was removed (the yellow bars), the two unorganized initial draft conditions led to final texts which were significantly better than the corresponding conditions where participants could retain their initial drafts. Crucially, these were equivalent in quality to those produced following the organized notes initial draft conditions. Removing the initial draft appeared, therefore, to enable the writers who produced initially unorganized drafts to revise these into well-organized texts.

Galbraith and Torrance's (2004) findings suggest a potentially effective revision drafting strategy would involve writers producing initial, synthetically planned drafts which could then be revised into effective final texts. To date, however, there has been relatively little other research into this possibility (with the exception of a study by Kieft, Rijlaarsdam, Galbraith, & van den Bergh, 2007, which investigated a similar, though not identical, revising strategy). The one exception is a recent study by Ong and Zhang (2013) examining the effectiveness of different drafting strategies for Chinese EFL learners writing in English.

The design of the study was similar to Galbraith and Torrance's (2004), and divided writing into three broad phases: (i) initial draft (30 minutes); (ii) summary & review (5 minutes); followed by (iii) a draft of the final text (20 minutes). Crucially, half the writers had the initial draft removed, while half were allowed to retain it, during phase 3. The initial draft conditions were rather different to those in Galbraith and Torrance's study, but included two planning conditions similar to outlining, and a control condition in which participants were instructed to write as normal. They also included a freewriting condition, similar to the unorganized sentences condition of Galbraith and Torrance's study. The findings are shown in Figure 3b. The key result is that, although participants wrote best in the control conditions, and surprisingly rather badly in the two planning conditions, removing the initial draft had a dramatic effect on performance in the freewriting condition. With the initial draft present, the participants produced the lowest quality texts in the whole experiment. By contrast, when the initial draft was removed, text quality was dramatically better (and was equivalent in quality to the control condition).

Taken together, these studies suggest that removing the initial draft – and hence *rewriting*, rather than *revising*, the initial draft – may enable writers to produce just as effective text following an unorganized initial draft as they can using a traditional outlining strategy or their own usual drafting strategy. Ong and Zhang's (2013) findings suggests that this applies in L2 just as much as in L1. The key claim of the dual-process model is that such a strategy would not only enable writers to produce better quality text but also to develop their understanding better during the initial draft than an outline-planning strategy would. Although this was not tested in these studies, it could be tested in future research by including a measure of subjective knowledge change. A further untested assumption of the dual-process model is that such a strategy, in enabling writers to better capture their implicit understanding of the topic, would make writing a more engaging process and hence increase motivation to write.

Implications for L2 writing research and language development

The dual-process model claims that knowledge (and text) are the joint product of two different processes – a knowledge-constituting process driven by the implicit structure of the writer's disposition towards the topic, and a knowledge-transforming process driven by external rhetorical goals. Moment-by-moment, any specific utterance is jointly determined by the writer's current goals – the individually represented social situation – and the current state of activation of the connections constituting the individual's disposition towards the topic – the writer's knowledge. Crucially, the strength of the connections within the writer's disposition are the product of an individual's learning history, including the social contexts within which the learning took place. Taken together, these features of the model correspond to many of the characteristics claimed by current theories of writing as social action (e.g. Bazerman, 2016; Cooper & Holzman, 1989; Prior, 2006; Russell, 1997). In particular, the model provides a specification of *voice* in writing (see Ivanič, 1998; Prior, 2001; Tardy, 2018) as the joint product of the writer's disposition (formally defined as the strength of the connections within a constraint satisfaction network), their linguistic knowledge, and their rhetorical goals. The writer gives voice to their thoughts not so much through dispassionate reflection on pre-established ideas followed by deliberate adaptation to the current communicative context but rather through acting in the current context on the basis of their accumulated experience and then reflecting on the outcome of their actions. Cycles of action and reflection enable the writer to gradually shape their thoughts into a rhetorically appropriate form. In this respect, the reconceptualization of the cognitive processes involved provides a way of integrating cognitive accounts of the writing process with the more sociocultural approaches described by Cumming (Chapter 2).

The account of text production as a knowledge-constituting process has strong similarities with Byrnes's (Chapter 4) characterization of writing as *textual meaning-making* and Byrnes and Manchón's (2014, p. 6) claim that writing "is about creating new textual worlds where language plays a constitutive role". Importantly, however, Baaijen and Galbraith's (2018) findings suggest that this is not an intrinsic feature of writing, but depends specifically on writing being synthetically rather than outline planned. Furthermore, it is not a single process: The full constitution of the writer's knowledge in the final text depends also on more reflective processing about the extent to which text satisfies rhetorical goals.

A key question here is how these two processes are affected by the writer's proficiency in L2. Perhaps, for example, the writer's ability to develop their understanding though spontaneous text production, and to develop their understanding further during revision, depends on their having reached a certain threshold of L2 proficiency. This is not immediately obvious though. The essential feature of the

knowledge-constituting process is that implicit thought is gradually constituted in a series of utterances, spontaneously produced in response to the preceding text. L2 proficiency would be expected to influence the efficiency with which writers produced individual utterances – less proficient writers would be expected to include less content in each utterance (Al-Saadi, 2018; Chenoweth & Hayes, 2003; Hayes, 2009) – but not necessarily to affect the production of successive dispositional responses to the emerging text. Similarly, the identification and reorganization of ideas in an initial draft may be less efficient for writers with low L2 proficiency, but may nevertheless still foster developments of understanding. It may be, therefore, that, even for writers with low L2 proficiency, a revision drafting strategy would enable writers to develop their implicit understanding of content more fully. Furthermore, even if it turned out that low L2 proficiency prevented writers from developing their understanding when writing in L2, it might nevertheless be valuable for them to write and revise an initial draft in L1 – to make their implicit understanding of the topic more explicit – before turning to producing further drafts in L2. These are ultimately empirical questions, which could be answered by a systematic programme of research assessing how L2 proficiency affects the development of understanding through writing, and the relative effectiveness of planning and revision strategies for writers with different levels of L2 proficiency.

Separate from these questions about the effect of L2 proficiency on writers' ability to constitute their understanding in text is the question of what the dual-process model implies about potential effects on language learning (as raised particularly by Cumming, Chapter 2). Our key claim here is that, in assuming that learners' knowledge is largely predetermined, and that writing is a matter of organizing and presenting this in an appropriate rhetorical form, classical problem-solving models of writing have neglected the knowledge-constituting potential of writing, and the opportunity this might offer for language learning.

There is evidence that this is not just a theoretical difference, but also distinguishes the conceptions that students have of writing, and that these affect the writing process. Baaijen, Galbraith and de Glopper (2014), for example, found that writers who have high transmissional beliefs (beliefs that writing is about transmitting pre-determined knowledge) produced higher quality text when they made an outline before writing than when they planned synthetically. By contrast, writers with high transactional beliefs (beliefs that writing involves developing thought in the course of writing) wrote equally well following either form of planning, but developed their understanding more than writers with high transmissional beliefs. Furthermore, for writers with high transactional beliefs, revision was associated with the development of their understanding, whereas for writers with low transactional beliefs, revision took a more surface form, and was directed at remedying linguistic problems with the text. Al-Saadi (2018) found, in a study of Omani writers, that

writers had stronger transmissional beliefs about L2 writing than L1 writing, and that beliefs became progressively less transmissional the greater the writers' L2 proficiency. These findings indicate that writers' beliefs about the writing process (White & Bruning, 2005) – or their metacognitive definitions of the process (Cumming, Chapter 2) – strongly influence how they go about writing and the effects that it has on their thought. We want to argue here that such differences in metacognitive beliefs may also affect the language learning potential of writing in L2.

If writers define writing as primarily a matter of translating pre-determined ideas into well-formed text, then revision is likely to be primarily directed towards correcting or modifying the way that the text is expressed. Although this opportunity to revise expression may provide some language learning potential (see Cumming, Chapter 2), we would argue that this potential will be enhanced by an expanded conception of writing as a knowledge-constituting process. One way of encouraging this would be to ask writers to write initial drafts of text directed towards capturing their understanding of the topic rather than trying to produce a well-formed and finished text. If, as Baaijen and Galbraith's (2018) results suggest, this enables them to further develop their understanding of the topic, and if, as the dual-process model (Galbraith, 1999; Galbraith & Baaijen, 2018) claims, this is because the knowledge-constituting process involves the progressive refinement of an initial synthesis of thought, then we would expect that, as initial formulations are inhibited, so alternative formulations should become more available (see Murphy & Roca de Larios, 2010). In effect, the attempt to capture implicit thought over a series of utterances would draw the writer into their zone of proximal development (see Cumming, Chapter 2), and enable them to arrive at a more precise representation of the content they were trying to express. In our experience, this sometimes occurs in the course of the knowledge-constituting process; on other occasions, however, it occurs later, during the revision process, when the writer tries to synthesize the range of alternative formulations into well-formed text. These observations could be explicitly tested by examining whether writing under such conditions enables writers to formulate a wider variety of linguistic expressions and, ultimately, arrive at a more precise expression of their thought, than writing under outline planned conditions does. It would be important also to test whether any such effects were moderated by the L2 proficiency of the writer.

Having discussed some of the potential implications of the dual-process model for our understanding of writing as a process of textual meaning making and the opportunities that writing provides for language learning, we want to conclude by considering some more methodological implications. The first of these is the importance of measuring the development of the writer's subjective understanding – this is the key indicator of the extent to which the writer has been able to capture their knowledge in the text. In taking this for granted, and treating writing as a matter of

adapting pre-existing ideas to rhetorical demands, problem-solving approaches to the study of writing have neglected the processes involved in capturing understanding in text. We view this as an important part of the writer's motivation in writing, so the fact that it is suppressed by the traditional outlining strategy may indicate that such strategies suppress the writer's ability to develop a voice in writing. We would argue, therefore, that research into writing should include explicit measures of the internal effect that writing has on the writer's understanding as well as measures of the external effect that it has on readers. In our current research, we use a multi-item scale to measure this, rather than the single rating used by Baaijen and Galbraith (2018), with the expectation that this will provide a more precise measure, capable of capturing different components of the writer's subjective knowledge.

The second implication is the need for research to investigate how writing in L2 affects the development of the writer's understanding. Although this is unexplored territory, one might expect from the dual-process model that the L2 would have differential effects on the different components of knowledge change, and that effects might vary depending on the writer's L2 proficiency. Does, for example, writing in L2 reduce the extent of the knowledge-constituting process because difficulties in formulating thought in language inhibit the forward progression of the writer's thought? By contrast, if the writer steps back from text production and focuses on organizing their thoughts when they are represented in an abbreviated linguistic form, are they as able to organize their thoughts in L2 as in L1? This would imply that, so long as the writer has an effective strategy for managing cognitive load, the knowledge-transforming component of the writing process should be similar in L2 and L1. These are important empirical questions for future process-oriented research interested in analyzing writing and language learning by writing and through writing (see especially chapters by Byrnes, Leow, Manchón & Leow, and the Coda chapter).

Third, Baaijen and Galbraith's (2018) empirical findings support the value of using keystroke logging to examine variations in how text production takes place during writing. (It is important to note here, however, that such measures are not transparent reflections of underlying cognitive processes [Baaijen et al., 2012]). Recent research, which combined keystroke logging with eye-tracking and / or verbal protocols to study L2 writing (e.g. Chukharev-Hudilainen, Saricaoglu, Torrance, & Feng, 2019; Révész, Michel, & Lee, 2019; López-Serrano, Roca de Larios, & Manchón, 2019; see also López-Serrano et al., Chapter 10, this volume), suggests some promising directions for developing more precise accounts of the processes involved and the language learning effects that may occur.

Several studies support the idea that sentence boundaries mark an important "hinge" in the writing process (Galbraith & Baaijen, 2019), with writers varying in how they carry out both within-sentence operations and recursive operations across the global structure of the text. The studies by Révész et al. (2019) and

Chukharev-Hudilainen et al. (2019) found that sentence boundaries were more likely to be associated with re-reading and revision of earlier sections of text, and that such re-reading and revision were more frequent in L2 than in L1 writing. They also found subtle differences in the form of sentence production in L1 and L2. Chukharev-Hudilainen et al. (2019), for example, found that pauses at clause boundaries were typically elevated in length compared to general pre-word pauses in L1, but that this difference wasn't present in L2, suggesting, perhaps, that sentence production was less hierarchically structured in L2 (Galbraith & Vedder, 2019). To date, however, such research has rarely examined the relationship between keystroke measures and text quality in L2, and no research has examined how these measures are interrelated with the development of the writer's understanding in L2. Baaijen and Galbraith's (2018) findings suggest that this is a key factor that needs to be taken into account. Are the variations that have been found in how writers produce sentences in L2 (Al-Saadi, 2018; Chukharev-Hudilainen et al., 2019; Révész et al., 2019) related to the extent to which they develop their understanding during writing, or do they reflect compensatory processes designed to ensure text quality in L2? Similarly, are higher levels of global revision in L2 associated with increased understanding (as in the synthetic planning condition studied by Baaijen and Galbraith, 2018), or do they reflect compensatory processes designed to maintain text quality, which are unrelated to developments of writers understanding (as in Baaijen and Galbraith's [2018] planning condition)? In order to answer such questions, the relationships between the more elaborate measures of online processing that have been developed in recent research and both text quality and knowledge change need to be established.

Finally, stepping back from the details of the processes involved in different components of the L2 writing process, the dual-process model raises questions about the effectiveness of different drafting strategies for L2 writers. In L1, the assumption is that writing a synthetically-planned initial draft, which is then reorganized and rewritten, enables the writer to first constitute their implicit knowledge in the text and then to present the text in a rhetorically appropriate form. However, it does not necessarily follow that such a drafting strategy would have the same effects in L2. For example, if writers are less able to constitute their implicit knowledge when writing in L2 than L1, then it may be more effective for writers to write an initial, knowledge-constituting draft in L1, and then rewrite this in L2, building on the clearer understanding developed in the initial draft. Such a strategy would be consistent with the assumption that the key feature of a multiple-drafting strategy is the opportunity it provides for the writer to capture their implicit knowledge of the topic, and with the tenets of the task repetition literature (see Chapters 6 and 7, this volume). Alternatively, it may be that the benefits that Ong and Zhang (2013) observed were a consequence of the opportunity to revise the language produced in

the initial draft, and that rewriting the draft is more effective than revising it. This would imply that the benefits of the task-repetition-like strategy would not be associated with developments of the writer's understanding but rather with improvements in language across the drafts and hence with potential language learning effects (see Chapter 1, this volume). In order to test these alternative possibilities, a series of studies are needed to establish how the writer's understanding and the linguistic features of the text develop across drafts, and how this is moderated by the language of the initial draft and the writer's L2 language proficiency.

Conclusion

Our fundamental claim in this chapter has been that, despite the ubiquity of the term "knowledge transforming" to characterize expert writing processes, the processes involved have not been specified fully and empirical research has not directly assessed effects of writing on the writer's subjective understanding. We have presented the dual-process model as a fuller specification of the processes involved, and illustrated the deeper understanding that this provides of the writing process by describing the results of recent research using keystroke logs. We suggest that future research in L2 designed to explore the connection between writing and learning needs (i) to incorporate explicit measures of the writer's subjective understanding; (ii) to use this to explore relationships with L2 language proficiency and with text quality; and (iii) to track how writer's understanding evolves over the course of writing.

References

Al-Saadi, Z. T. (2018). The effects of proficiency in English as a Foreign Language, gender and writing beliefs on the writing processes and products of undergraduate Omani students (Unpublished doctoral dissertation). University of Southampton.

Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, 36(3), 199–223. https://doi.org/10.1080/07370008.2018.1456431

Baaijen, V. M., Galbraith, D., & de Glopper, K. (2014). Effects of writing beliefs and planning on writing performance. *Learning and Instruction*, 33, 81–91. https://doi.org/10.1016/j.learninstruc.2014.04.001

Baaijen, V. M., Galbraith, D., & de Glopper, K. D. (2012). Keystroke analysis: Reflections on procedures and measures. *Written Communication*, 29, 246–277. https://doi.org/10.1177/0741088312451108

Bazerman, C. (2016). What do sociocultural theories of writing tell us about learning to write? In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 11–23). New York, NY: The Guilford Press.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Byrnes, H., & Manchón, R. M. (2014). Task, task performance, and writing development: Advancing the constructs and the research agenda. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 267–299). Amsterdam: John Benjamins.

Chenoweth, N. A., & Hayes, J. R. (2003). The inner voice writing. *Written Communication*, 20(1), 99–118. https://doi.org/10.1177/0741088303253572

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H.-H. (2019). Combined deployable keystroke logging and eye-tracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, 41, 583–604. https://doi.org/10.1017/S027226311900007X

Cooper, M. M., & Holzman, M. (1989). *Writing as social action*. Portsmouth, NH: Boynton/ Cook.

Cumming, A. (2016). Theoretical orientations in L2 writing. In R. M. Manchón & Matsuda, P. K. (Eds.), *Handbook of second language and foreign language writing* (pp. 65–88). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-006

Fitzgerald, J. (2006). Multilingual writing in preschool through 12th grade: The last 15 years. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 337–354). New York, NY: The Guilford Press.

Flower, L. S. (1979). Writer-based prose: A cognitive basis for problems in writing. *College English*, 41(1), 19–37. https://doi.org/10.2307/376357

Flower, L. S., & Hayes, J. R. (1980a). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, 31(1), 21–32. https://doi.org/10.2307/356630

Flower, L. S., & Hayes, J. R. (1980b). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive processes in writing: An interdisciplinary approach (pp. 31–50). Hillsdale, NJ: Lawrence Erlbaum Associates.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 139–160). Amsterdam: Amsterdam University Press.

Galbraith, D. (2009). Writing as discovery. *British Journal of Educational Psychology*, 2(6), 5–26. https://doi.org/10.1348/978185409X421129

Galbraith, D., & Baaijen, V. M. (2018). The work of writing: Raiding the inarticulate. *Educational Psychologist*, 53, 4, 238–257. https://doi.org/10.1080/00461520.2018.1505515

Galbraith, D., & Baaijen, V. M. (2019). Aligning keystroke with cognitive processes in writing. In E. Lindgren & K. P. H. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 306–325). Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004392526_015

Galbraith, D., & Torrance, M. (2004). Revision in the context of different drafting strategies. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision of written language: Cognitive and instructional processes* (pp. 63–85). Dordrecht, The Netherlands: Kluwer. https://doi.org/10.1007/978-94-007-1048-1_5

Galbraith, D., & Vedder, I. (2019). Methodological advances in investigating L2 writing processes: Challenges and perspectives. *Studies in Second Language Acquisition*, 41, 633–645. https://doi.org/10.1017/S0272263119000366

Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187–207). New York, NY: Guilford Press.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy, & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, J. R. (2009). From idea to text. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *The Sage handbook of writing development* (pp. 65–79). London: Sage. https://doi.org/10.4135/9780857021069.n5

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. https://doi.org/10.1177/0741088312451260

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American Psychologist*, 41(10), 1106–1113. https://doi.org/10.1037/0003066X.41.10.1106

Ivanič, R. (1998). *Writing and identity: The discoursal construction of identity in academic writing.* Amsterdam: John Benjamins. https://doi.org/10.1075/swll.5

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2), 355–365. https://doi.org/10.1037//0278-7393.14.2.355

Kellogg, R. T. (1990). Effectiveness of prewriting strategies as a function of task demands. *American Journal of Psychology*, 103, 327–342. https://doi.org/10.2307/1423213

Kellogg, R. T. (1994). *The psychology of writing.* Oxford: OUP.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research*, 5, 159–90. https://doi.org/10.17239/jowr-2013.05.02.1

Kieft, M., Rijlaarsdam, G., Galbraith, D., & van den Bergh, H. (2007). The effects of adapting a writing course to students' writing strategies. *British Journal of Educational Psychology*. 77, 565–578. https://doi.org/10.1348/096317906X120231

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, 20(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30, 358–392. https://doi.org/10.1177/0741088313491692

Lindgren, E., & Sullivan, K. P. H. (Eds.) (2019). *Observing writing: Insights from keystroke logging and handwriting.* Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004392526

López-Serrano, S., Roca de Larios, J., & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks. *Studies in Second Language Acquisition*, 41, 503–527. https://doi.org/10.1017/S0272263119000275

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35(1), 13–23. https://doi.org/10.1207/S15326985EP3501_3

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

Murphy, L., & Roca de Larios, J. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL writers. *Journal of Second Language Writing*, 19(2), 61–81. https://doi.org/10.1016/j.jslw.2010.02.001

Ong, J., & Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers' text quality. *TESOL Quarterly*, 47 (2), 375–398. https://doi.org/10.1002/tesq.55

O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science*, 38(6),1229–1248. https://doi.org/10.1111/j.1551-6709.2011.01214.x

Prior, P. (2001). Voices in text, mind, and society: Sociohistoric accounts of discourse acquisition and use. *Journal of Second Language Writing*, 10, 55–81. https://doi.org/10.1016/S1060-3743(00)00037-0

Prior, P. (2006). A sociocultural theory of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54–66). New York, NY: The Guilford Press.

Révész, A., & Michel, M. (2019). Methodological advances in investigating L2 writing processes: Introduction. *Studies in Second Language Acquisition*, 41(3), 491–501. https://doi.org/10.1017/S0272263119000329

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors: A mixed-methods study. *Studies in Second Language Acquisition*, 41, 605–631. https://doi.org/10.1017/S027226311900024X

Roca de Larios, J., Nicolás-Conesa, F., & Coyle, Y. (2018). Focus on writers: Processes and strategies. In R. M. Manchón & Matsuda, P. K. (Eds.), *Handbook of second language and foreign language writing* (pp. 267–286). Berlin: De Gruyter Mouton.

Russell, D. R. (1997). Writing and genre in higher education and workplaces. *Mind, Culture and Activity*, 4(4), 224–237. https://doi.org/10.1207/s15327884mca0404_2

Stevenson, M., Schoonen, R., & de Glopper, K. D. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, 15, 201–233. https://doi.org/10.1016/j.jslw.2006.06.002

Tardy, C. (2018). Voice and identity. In R. M. Manchón & Matsuda, P. K. (Eds.), *Handbook of second language and foreign language writing* (pp. 349–363). Berlin: De Gruyter Mouton.

White, M. J., & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology*, 30(2), 166–189. https://doi.org/10.1016/j.cedpsych.2004.07.002

Toward an agenda for researching L2 writing and language learning

The educational context of development

Heidi Byrnes
Georgetown University

Taking an educational linguistics perspective, the chapter highlights four areas that have the potential of illuminating possible effects of L2 writing for L2 learning. They address pressing issues by (a) theorizing writing as textual meaning-making; (b) emphasizing the pivotal role of the educational context, in its conceptualization and its practice; (c) affirming the long-term nature of developing literate forms of language use under a longitudinal optic; and (d) privileging an extended, well-conceptualized curricular framework. The chapter suggests that these dimensions can yield important insights regarding the long-term development of both L2 composing and L2 abilities, especially when they are informed by a textually oriented theory of language, such as Systemic Functional Linguistics, and contemporary thinking in Complex Dynamic Systems Theory.

Introduction

In this chapter I identify areas of inquiry that I consider to be particularly promising for advancing our understanding of the interrelationship between L2 writing and language learning. Foundational for my reflections is a critical examination of instructed L2 writing as an ontological phenomenon. This step is crucial because L2 writing research has over time espoused theoretical, methodological, and research-empirical epistemologies that differ with respect to viewing language use in the written mode as a particular *form of communication,* a characterization I consider to be criterial. Specifically, research has evolved toward privileging a conceptualization of the writing-language learning interface in terms of cognitive processes and their manipulation, a decision that has affected both how it imagines and studies the act of composing itself and how it addresses learners' subsequent engagement with corrective feedback on their compositions. Strongly paired with

a focus on sentence-level phenomena of linguistic form – primarily analyzed in terms of complexity, accuracy, and fluency of performance – this preference has resulted in powerful metaphors, analogies, and practices in L2 writing research. These prevail even when other non-processing oriented theoretical frameworks, such as Sociocultural Theory, Skill Acquisition Theory, or usage-based approaches, are invoked and when educational considerations are explicitly affirmed (see Leow, Chapter 5, and Manchón & Leow, Chapter 14, this volume; Leow, 2019; Leow & Cerezo, 2016).

The task of envisioning future research agendas for L2 writing-language learning connections thus faces the following challenges: First, to appraise the epistemological thrust, intellectual significance, and educational value of findings that we currently hold as sedimented knowledge; second, to assess these findings in light of well-known impasses, persistent empirical incongruities, and inconclusive, even contradictory, recommendations for how educational programs might foster language learning through writing; and third, to consider the extent to which lacunae in our knowledge regarding the writing-language learning interface might reflect conceptual and empirical habits of mind that by now merit careful re-evaluation.

Embarking on this kind of inquiry calls for taking a phenomenological-ontological stance toward L2 writing in order to mitigate the undue influence of theoretically driven assumptions. It might ask questions such as these: What are the core non-negotiable characteristics of writing as a form of engagement with language that distinguish it from the other modalities? What aspects of writing as a particular form of language use – whether in the L1 or additional languages – cannot be compromised because they constitute its very own, perhaps even unique qualities? In short, what aspects of writing should constitute privileged forms of inquiry into and reflection regarding writing no matter where, with whom, and in what setting? At the same time, how do differences among the modalities nevertheless reflect and, indeed, facilitate, mutually supportive interrelationships between them, especially with reading, but also with speaking and listening? What particular challenges confront the L2 writer beyond those that are inherent to any writing, where a yet evolving ability to draw on a rich repertoire of L2-derived meaning-making resources is, of course, crucial? What aspects are likely to be particularly insightful with regard to imagining, understanding, and easing the long road to both L2 writing and L2 language development? And, finally, what benefits might we expect from research that would seek to answer questions such as these for the flourishing of individuals and societies in the age of multilingualism, globalization, and digital forms of communication?

I will pursue these queries in four interrelated areas: (a) theorizing writing as textual meaning-making that, because of its elaborated and expansive form of conveying intended meanings, also fosters continued elaborated L2 language learning;

(b) foregrounding the pivotal role of the educational context for shaping both L2 writing and L2 language learning; (c) adopting a theoretically principled longitudinal optic that is learning-developmental rather than merely chronological; and (d) privileging a well-conceptualized curricular framework in order to arrive at educationally useful and usable statements and recommendations for fostering writing and language use abilities. Because understanding writing as textual meaning-making poses the greatest conceptual-theoretical challenge I will treat it more expansively. At the same time, the most far-reaching conceptual-empirical-practical challenges are likely to arise from the remaining three topics, all of which relate to how the educational setting can cultivate L2 composing and L2 language learning abilities.

Theorizing writing as textual meaning-making

I propose as *the* overriding focus for theorizing writing its quality of *textual* meaning-making. That proposal reprises arguments offered by Byrnes and Manchón (2014a, p. 6) in the context of task-based teaching: "No matter what else composing is and does, it is about creating new textual worlds where language plays a constitutive role, even in an increasingly multimodal world." Two aspects of this definition speak directly to a potential interface between writing and L2 learning: The *textual expansiveness* of writing that, typically, is realized by a single authorial voice, as contrasted with the relatively short, distributed, and interactively negotiated turn-taking nature of spoken language; and the *heightened functional load* that language itself must take on.

It is worth recalling that construal of a social-situational context is necessary for any meaning-making, no matter the modality. In oral communication that requirement is facilitated by the setting of the interaction itself. However, in written texts, the interpretive context has to be created with language. In its fully fledged manifestation this literate language must be able to realize an elaborated dialogic discourse between writer and imagined reader(s) that *expresses in and with language* what oral language accomplishes with the support of the physical setting in direct interactive exchange.

I take both the potential for expansive elaborateness of writing and the greater functional demands being placed on language to be criterial for any writing. While such a characterization may not describe the bulk of L2 writing in instructional contexts and seems best reserved for more advanced performance levels, I reject that restricted interpretation: Not only are even beginning-level writers at all educational levels quite capable of considering texts as a mode of communication that differs from oral language use and of adjusting their language resource use accordingly; more consequential, not allowing for that capacity early on seriously

undermines the entire argument for an L2-writing-language learning link with obvious consequences for investigating it substantively.

Two additional misconceptions must be addressed. First, greater functional demand for literate language is not the same thing as greater complexity, whether syntactic or lexical; at the very least, it requires us to understand 'complexity' in ways that go beyond "more is better" assumptions. And second, it is the textual qualities of written language, more than additional processing time, that strongly supports the proposition that writing facilitates language learning. This is so because under an optic that observes L2 writers gradually appropriating a rich palette of textually oriented semiotic resources beyond those that characterize oral modes of communication it becomes possible to trace their impact on language learning over an extended period of time.

In the remainder of this section I endeavor to make that case from three perspectives. I begin by relating Cumming's seminal work in the 1990s to present-day cutting-edge psycholinguistic processing-oriented research in order to explore the possibility of common ground between a processing and a meaning orientation. I then consider two conceptual treatments that highlight the uniqueness of writing, by Emig and Vygotsky, and provide a close-up view of composing offered by Witte that suggests promising approaches to researching the act of composing as a meaning-making activity that fosters language learning. I conclude the section by exploring how imagining the L2 writer as textual meaning maker might advantageously unite previously separate research directions.

Linking processing and meaning-oriented approaches to compositional writing

At the outset I argued that textual meaning-making is a defining characteristic of any writing. For readers familiar with the tasks that serve as the data sources for much L2 writing research, such a statement may appear to be wishful thinking. As frequently noted, too many composing prompts are contrived, repetitive, unimaginative, and devoid of the potential of thoughtfully engaging writers. Rather than belaboring and bemoaning that fact, what matters instead is to ascertain just how much and what kind of meaning-making L2 writers actually do engage in. Remarkably, few such studies exist, which makes Cumming's (1990) study designed to describe the decision-making processes participants used in the act of composing all the more noteworthy.

Specifically, he found considerable metalinguistic and ideational thinking in L2 composing, particularly when L2 writers are asked to engage with writing tasks that enable them to "believe that the substance of their writing merits careful thought,

that the purpose of writing is to convey information to others, and the texts they produce can be improved through rethinking and revision" (p. 504). In that case, verbal protocols about their writing show them to exercise "a high degree of intentionality" where decision-making is often *metalinguistic and ideational* [original italic]… [in that] learners focus mental effort directly on language items while concurrently thinking about their ideas and semantic meanings" (p. 500) as they search for and assess appropriate words and phrases, compare crosslinguistic equivalents, and reason about linguistic choices.

The fact that the latter strategy, which relies primarily on the knowledge of linguistic rules in order to convey an intended meaning, occurred only rarely in Cumming's data invites critical reflection. First, it serves to underscore that verbalized reports pertaining to decisions regarding language form are not the same thing as being preoccupied with rule-based searches for accurate language form nor do they preclude an overall meaning-orientation of compositional writing. This is the provocative message behind Cumming's statement that "these thinking processes are *incidental* to goals of effective communication" (p. 504).

Second, invoking as the *primary reason* for writers' searching behavior their emerging and therefore incomplete repertoire of L2 resources deserves considerable caution. Such an interpretation can easily stray into deficitary views of L2 writers, not only with regard to their ability to process the language 'fluently' or possessing formal and preferably 'accurate' writing skills. More consequentially, it brands them as fundamentally impaired meaning-makers, thereby occluding, right from the beginning, exactly what we wish to understand, namely the L2 composing – language learning interface as it manifests itself in the *written mode of communication.* To state the obvious: There is little gain in restating that the L2 writer's command of the semiotic resources of the language she is learning is incomplete. What is considerably more interesting is that for L2 writers, just like for native writers, "intended meaning" is dynamically emergent and, crucially important, language-based. In other words, intended meaning itself evolves in the particular form of language use we call writing precisely because of the opportunity for multilevel and nonlinear reflection on diverse meaning-form interfaces that it affords.

Third, a psycholinguistic orientation, particularly when it is interpreted as "a psycholinguistic output condition" (p. 483) does not exclude other interpretive stances, especially when recent studies have increasingly distanced themselves from narrowly conceived notions of processing (see Polio, Chapter 16, this volume) and have come to frame in a much more differentiated and rigorous fashion how problem solving, hypothesis testing, and depth of processing (DoP), including writers' engagement with corrective feedback, ultimately points to the writers' underlying goal of realizing their intended meanings in context. As Craik (2002, p. 309) portrays

the issue with regard to DoP: "Any valid index of depth must therefore measure the *meaningfulness* and elaboration of the final encoded representation, and not simply the ease or difficulty of achieving that representation" (emphasis added).

For example and representative of such recent work, the excellent discussion in López Serrano, Roca de Larios, & Manchón, (Chapter 10, this volume) encourages us to consider "the combined effects of task-type and proficiency in promoting DoP and bringing about language learning through writing" (p. 247). To me, task-type is a proxy for the communicative or meaning-oriented intent one can expect a text to aim at. It is that complex interrelationship between intended meaning and linguistic resources that affects learners' orientation either toward compensatory changes due to "writers' lack of access to linguistic knowledge required to express their intended meaning" or, alternatively, to an upgrading orientation that seeks to "improve the expression of one's intended meaning" (p. 238). One should not be surprised that, among other factors, processing orientation is strongly related to proficiency level. More revealing is that the researchers, as they invoke Hulstijn's (2011) conceptualization of L2 proficiency in terms of basic language cognition and higher language cognition, enjoin us to consider carefully both the effects of L2 proficiency and the differential effects of DoP *in relation to the writing tasks* that the study's participants were asked to handle. Their conclusion that an argumentative task might not be most appropriate "to take full advantage of the language learning potential of writing" (p. 247) for intermediate-level learners is an indirect way of acknowledging that the underlying issue is not processing itself, but the kind of textual meaning-making that L2 writers are asked to engage in at different points of their long journey to becoming competent writers. In other words, the increasingly sophisticated awareness these writers have regarding the functional – that is, *the communicative* – consequences of deploying certain lexicogrammatical resources of their L2 must be of central concern to researchers.

To sum up, the best contemporary research seeking to illuminate the writing-language learning interface creates an advantageous space for imagining what Byrnes and Manchón (2014b, p. 271) envisioned as the potential for "interweaving both psycholinguistic processes … with intense linguistic activity associated with the meaning-making effort inherent to the very act of writing." Assuming that it continues to be the case that "very little is known about how thinking is linked moment by moment with the production of text itself or with the externally represented text as it is produced" (Roca de Larios, Nicolás-Conesa, & Coyle, 2016, p. 280), a treasure trove of future research opportunities opens up.

In search of the uniqueness of writing as embodied thinking with language: Emig and Vygotsky

Returning to the overriding question animating this segment of the chapter, I summarize an article written by Emig (1977) a little over four decades ago, where she asked this pivotal question with regard to native-language composing: What about writing is "not merely valuable, not merely special, but unique" (p. 122) as a *mode of learning*? Answering that question, she described writing as a second-order process (as contrasted with the first-order process of speaking) and one which typically requires instruction. She further characterized it as a dynamic "languaging process" that must be understood as both "originating and creating a unique verbal construct that is graphically recorded" (p. 123). To her, it is that combination of qualities that requires particular care inasmuch as it separates writing from speaking; granted, speaking also creates and originates a verbal construct, but it is one "that is *not* graphically recorded" (p. 123, original emphasis) and that is not embodied.

Accordingly, she specified the uniqueness of writing in terms of its near simultaneous and three-directional "symbolic transformation of experience through the specific symbol system of verbal language … shaped into an icon (the graphic product) by the enactive hand" (p. 124). In that fashion writing incorporates Bruner's "three major ways in which we represent and deal with actuality: (1) enactive – we learn 'by doing'; (2) iconic – we learn 'by depiction in an image'; and (3) representational or symbolic – we learn 'by restatement in words'" (p. 124). To that must be added the power of re-enforcement that arises from the multi-representational quality of writing; the active participation of all our mental capacities, most especially emotional commitments to what is said by attending to how it is said; the role of intuitions that first appear as visual and spatial wholes and energize the creative process; the "unique form of feedback … because information from the *process* is immediately and visibly available as that portion of the *product* already written" (p. 125); and the opportunity to connect the major times of our experience – past, present, and future – because of its slowed-down pace, and the fundamentally epigenetic quality of writing.

For contemporary readers it is striking that Emig powerfully positions L2 writing as a particular form of embodied thinking. Referring to the Russian psychologist Luria and his observation that speech, as contrasted with writing, involves the *synpraxis*, that is, the intricate interdependence of an actual situation and language use, she offered this quote by him to capture the unique nature of writing:

> Written speech is bound up with the *inhibition* of immediate synpractical connections. It assumes a much slower, repeated mediating process of analysis and synthesis, which makes it possible not only to develop the required thought, but even

to revert to its earlier stages, thus transforming the sequential chain of connections in a simultaneous, self-reviewing structure. Written speech thus represents a new and powerful instrument of thought.   (p. 127, emphasis added)

Reference such as this to a kind of breaking open in the process of writing what in oral language use is the close synpractical connection between the occasion for meaning-making and language use invites consideration of the work of a second scholar, the Russian psychologist Vygotsky (1986), who reflected extensively on thought and language in both speaking and writing. But before citing Vygotsky, I note that the interpretation provided by prominent writing researchers Swain and Lapkin in their 1995 study, itself a partial replication study of Cumming (1990), has provided a particularly influential metaphor for writing researchers. Specifically, it examined the claim that "'pushing' learners beyond their current performance level can lead to enhanced performance" (p. 374) as they encounter linguistic problems in the L2. In an oft-repeated formulation, the researchers interpreted the internal thought processes observed in the think-alouds as moving learners "from semantic to grammatical processing" (pp. 386, 388) in order to express their intended meaning. Such a formulation strongly suggests a pre-existent and stable meaning that is essentially independent of its realization in a linguistic form, whether in the L1 or the L2, and, for that reason, interprets the learner's task as one of finding and matching the correct L2 form to that pre-existing meaning.

Turning now to Vygotsky's reflections on the matter, he concluded that the grammar of thought differs in the case of speaking and writing. While the former draws on "maximally compact inner speech … maximally detailed written speech requires what might be called deliberate semantics – deliberate structuring of the web of meaning" (p. 182). In other words, the very activity of composing – as crystallized in the act of composing – occasions and indeed requires more deliberate *ways of meaning* that are reminiscent of Luria's positing of transformative new connections being made. Leaving open the precise ways in which this multidimensional connecting might take place, such a proposal establishes thought and meaning as being in flux in complexly relational ways. That fluidity is the result of the writer tentatively "holding in mind" the diverse semiotic functions being performed by particular lexicogrammatical options as she considers their suitability, desirability, and appropriateness for her particular composition. Writers' verbal protocols frequently portray this as a "searching for words" in order to express what they "had in mind all along"; but careful inspection of the protocols reveals a considerably more dynamic process involving *both* meaning and form and, importantly, involving much more – and much more diverse – linguistic material than words (see Schmitt, Chapter 15, this volume).

There is every reason to consider such fluidity in language-based meaning-making to apply to both L1 and L2 composing, though, quite understandably, in the latter case what linguistic resources are available to the writer deserves particular attention at the intersection of task and proficiency (see the earlier discussion of that point; also Macaro, 2014). As for implications for a research agenda, *how* L2 writers go about settling on the details of their language choices, *what* those choices are *with regard to their meaning-making function in a textual environment,* and *how they change and develop* over the course of the writers' evolving textual abilities offers a nearly unique window for the topic of this volume, the link between writing and language learning (for a glimpse into this potential, see, e.g., the examples provided in López Serrano et al., this volume; Macaro, 2014; Witte, 1987).

And an additional issue deserves attention: What is ultimately 'put down on paper' and what *was* 'meant' – and *could be* meant at the time – is not final, but what Complex Dynamic Systems Theory (CDST) refers to as "soft assembly" (e.g., Larsen-Freeman, 2011) in order to highlight the continuing emergence of capacities for language use. To paraphrase Halliday (1999), meaning-making *with* and *in* and *through* a language is a lifelong work in progress. At the same time, a cautionary interpretation both of what we call problem-solving, form-oriented processing, and what was 'meant' by the writer in a particular instance of writing provides us with a powerful, theoretically and empirically substantive reason for longitudinal study. Rather than being *negatively motivated* as a way of avoiding the "we did not study something long enough, therefore obtained inconclusive results" verdict of many a study, longitudinal study now has the *positive charge* of enabling us to trace the gradually changing meaning-making behaviors and patterns in written language use in terms of their complex heterarchical interrelationships between meaning and form.

Toward a textual meaning-oriented inquiry for L2 composing

Just that possibility, along with research methodological recommendations, is my take-away of a study by Witte (1987), in which he examined close-up the think-aloud data obtained by native composers. What stands out is his focus on the pre-text phase of composing prior to transcription, an analysis that leads him to conclude that it constitutes a critical component of any theoretical model for writing because it has "both a syntactic and a semantic component" (p. 397). In other words, not only are meaning and language form interpellated right from the start, but restricting revision to the retranscription of text that is already written down severely misconstrues the dynamic nature of language-based meaning-making. By extension, not only should we consider that "translating ideas into verbal language may be a more complex and variable composing process than has to date

been recognized" (p. 217); we may have to reconsider the extent to which planning and translating can at all be as neatly distinguished as has usually been assumed. Furthermore, such pre-text work in composing, what Witte refers to as "trial locution," for some writers extends not only beyond words, but far beyond a single sentence into constraints and affordances of an expansive textual discourse gauged in light of a particular situational context for a particular composition.

Insights like this facilitate a way of operationalizing 'learning' and 'development' in compositional language use differentially at different proficiency levels. At the same time, such an operationalization is always holistic, not separatist, inasmuch as it focuses on the L2 writer's understanding of the meaning-making consequences afforded by different lexicogrammatical options that she is considering. It is, then, not 'words' – the expression frequently used by lay writers – nor is it the complexity, accuracy, or fluency of their features and how they are processed – the focus of many researchers – that are of interest: Rather, of interest are the *choices* being made – and able to be made – by the L2 composer at the phrase, clause, sentence, complex sentence, discourse themes and hypertheme levels, all the way to the communicative goal and purpose of an entire text in social context, understood as a semantic unit. Because these choices play themselves out differently in composing different genres at different L2 proficiency levels, they offer a unique window into understanding the interrelationship between writing and language learning. It is worth quoting Witte's conclusion at some length:

> In my judgment, too much research on composing […] has insisted on a separation of process and product, thereby underestimating the importance of emerging text. What is very much needed is research […] which explores connections between composing processes and the emergence of written products. Such research would likely lead to a better understanding of how, during composing, writers attend to (or fail to attend to) the three simultaneous functions Halliday (1973, 1978) associates with language use: the ideational, the textual, and the interpersonal.
>
> (p. 418)

Toward researching the L2 writer as textual meaning-maker

Throughout this section I have offered glimpses for how reconceptualizing writing as textual meaning-making might influence a research agenda at the L2 writing – language learning interface. In an interim assessment of the argument thus far, I would like to suggest that three highly productive research areas – the learning to write, writing to learn, and writing to learn language perspectives – can and should now be conceptually linked (see the incisive treatment of these matters in Manchón, 2009, 2011; Manchón, Roca de Larios, & Murphy, 2009; Roca de Larios, 2013). Such a move would overcome the intellectual awkwardness and practical

challenge of attempting to research an overarching hypothesis in environments that in educational programs taken in their entirety are not only difficult to separate (but see Manchón & Leow, Chapter 14, this volume) but right up front truncate the evidence we need for its examination.

Fortunately, a more unified approach is already on the horizon. For example, earlier on I referred to Vygotsky's imagery of a web of meaning which, with a strong CDST stance, is reminiscent of Cumming's (2016) notion of "an ongoing mental dialectic between content and rhetorical concerns" (p. 70). Taking an explicitly processing oriented view, Galbraith and Al-Saadi (Chapter 3, this volume) proposes a dual-process model for writing, "in which writing is the joint product of two conflicting processes: An explicit problem-solving process … and an implicitly controlled knowledge constituting process, taking place during the formulation of thought in language, and responsible for developing the writer's understanding of a topic" (p. 50). Coming from a functional textual meaning-making side grounded in Systemic Functional Linguistics (SFL), Ryshina-Pankova and Byrnes (2013) offer *linguistic evidence* for how writers go about creating knowledge by using language resources that are particular to literate language, generally taken to represent advanced forms of language use.

In and of itself, such evidence from finished compositions cannot offer proof that writing *caused* language learning to that level. However, if such literate language use is consistently attained in an L2 educational setting that deliberately fosters it within its entire program, then it should yield insights regarding conditions that are favorable to the hypothesized relationship. The next section examines how an educational context might differentially illuminate such linkages.

Foregrounding the pivotal role of the educational context

The need to situate the writing-language learning connection within an instructed SLA perspective is a central motivation underlying this volume (see particularly, Leow, Chapter 5; and Manchón & Leow, Chapter 14). Hence, the challenge to the research community is this: Whenever possible, work collaboratively ahead of time with a program's educators to identify those issues for which they seek more differentiated answers *in line with major educational choices* made and implemented in their program. Such a collaboration is meaningful for researchers and educators alike even in settings where curricular guidelines are government-mandated and, therefore, relatively fixed. Joint reflective practice will make both parties aware of myriads of factors that can come into play and that may have outsized consequences for the kind of learning that takes place. Furthermore, each of them individually and in ever changing configurations will complexly interact with the others, – in

different settings, for different learners, at different stages of their writing career, at different time scales, and different levels of observational granularity. Taken together, these factors further bolster the validity of imagining a research enterprise that (a) takes seriously the intricately interwoven *educational* nature of writing development, and yet proposes a reasoned way of implementing a coherent writing program; (b) privileges its variable longitudinal manifestations over extended instructional sequences; and (c) observes it within specific settings.

Exploring educational dimensions of writing development

All educational settings reflect choices – and that means diverse power relationships and values orientations exclude some options while privileging others. Not surprisingly, those choices are subject to periodic public scrutiny and changes in educational policy, with the development of literacy in a second language or the failure to do so frequently becoming particularly contentious territory. At the same time, it is precisely the reduction of complexity in particular educational settings (see Biesta, 2010) that results in the possibility of researching educational programs in a way that facilitates a better understanding of change and development in that setting.

Specifically, our understanding of development is significantly enhanced if we relate it to the constitution of an entire system with its dynamic interrelationships among components that have been identified as constitutive of that system. Yet more precisely, it is emergence, "in which the interactions among components both with each other and with the whole of which they are part are constitutive of properties of the systems" (Byrne & Callaghan, 2014, p. 22), that is at the core of change, and therefore of any possible development. And it is emergence rather than fixed goals that "*opens up multiple new possibilities* (Osberg, 2008, p. 157, original emphasis) for educational renewal that is suited to particular settings.

What might this abstract formulation, which articulates the system view of social reality espoused by CDST, mean in the context of education that would allow us to consider "interactions among components … and with the whole" in order to observe the emergence of learning that, eventually, might turn into development? I suggest the following dimensions as potential candidates for fostering that kind of emergent writing and language learning and, with appropriate research design, allowing us to trace it:

- Separate and/or integrated development of writing within all modalities;
- Explicit and/or incidental incorporation of writing;
- Writing from the beginning and/or delay until a certain specified proficiency level;

- Primarily in-class or out-of-class writing with and/or without specified access to support, such as dictionary or grammar use or access to diverse instructional materials;
- Overt linking of reading and writing in order to foster literacy development;
- Integration of writing into content-based approaches to language teaching, especially with an emphasis on source-based writing which inherently involves reading;
- Different weighting of composing assignments and/or tasks that require the use of textual sources.

Given the *central role of genres* in some instructional settings, additional guiding questions might be:

- Does the program instantiate a progression of the most important macro-genres (e.g., narrative, explanatory, argumentative) and how are students made aware of that progression?
- What assumptions about the nature of writing development in relation to language learning and knowledge construction in the disciplines and their central genres animate the program?
- What counts as progress in writing ability in a genre-based approach, and how is that conveyed to the students?

Only once such broad program choices have been made and implemented does it make sense – not least from a research cost-benefit perspective – to ask questions about specific pedagogies, both in terms of their highly contextualized nature *and* the reality of variable learning outcomes. For example, assuming that a program takes a genre-oriented approach to writing, we would want to know: How are lexicogrammatical features taught as aspects of textual meaning-making? How is the long-term nature of writing development conveyed pedagogically, particularly in terms of differential feedback practices and assessment of writing performance? How do pedagogical emphases differ at different performance levels with different genres and why?

It is this dual orientation – toward overriding programmatic features on the one hand and their embedded, contextually motivated pedagogical actions on the other – that provides fertile ground for desirable research projects that can illuminate the writing-language learning interface.

Affirming the long-term nature of developing written literacy

Calling for longitudinal research to capture development has become nearly normative in practically all domains of applied linguistics. What remains unclear, however, is how we are to imagine *development* in L2 writing and L2 language use in its extended evolution, under what configurations of educational practices for what learners with what realistic learning outcomes for what languages.

For longitudinal research to be able to yield answers to these kinds of questions we need well-considered proposals prepared by the educators of a particular program regarding: (a) dimensions of L2 writing development as a special mode of language use, what I have called its ontological quality, translated into possible phases/stages/levels in the program; (b) their understanding of texts and their language use and ways of fostering the development of literate language resources; and (c), their shared understanding of supportive pedagogical practices for both L2 writing with L2 language development across an extended period of time and with variable trajectories by individual students and entire student cohorts. To be maximally useful, such considerations should encompass issues arising from beginning to advanced levels of use. Only then can we begin to formulate worthwhile longitudinal research questions and determine their specifics like "the appropriate time scale, duration, and frequency of observations … in researchable terms" (Norris & Manchón, 2012, p. 228). Only then can we hope to obtain insightful and educationally actionable findings.

Among possible directions for longitudinal inquiry into writing development I suggest the following:

First, directing Witte's processing-oriented approaches toward the act of composing at the pre-text stage could affect both the scope of inquiry – that is, the expansiveness of the linguistic units being 'held in mind' by the writer – and the function of those units in terms of their rhetorical effectiveness as the writer perceives it. Importantly, it would frontload the L2 writing research agenda in contrast with its current backloading preoccupation with written corrective feedback (WCF) and its efficacy.

A second suggestion for orienting longitudinal inquiry can be found in cross-linguistic research as described in Berman and Verhoeven (2002). In studying how school children construct monologic texts, with a focus on the linguistic, cognitive, and communicative resources they deploy in narrative and expository texts, the study's key findings provide valuable guidance for a longitudinal L2 writing and language learning research agenda. For example, "the ability to recruit linguistic forms in the context of extended discourse … has a long developmental history" (p. 14) and forcefully reiterates the need for longitudinalness counted in years rather than

weeks and months. The researchers emphasize the need to "analyse linguistic forms in relation to their discourse function beyond the structuralist limits of isolated sentences, on the one hand, and of grammatical correctness, on the other" in order to capture phenomena of text structuring, thematic content, propositional attitudes and discourse stance (p. 15). Third, one should expect a "confluence of cues" (p. 15) as contributing to the expression of different discourse functions, such as a text's temporal texture or the discourse stance taken by the writer. Lexicon, rightly seen as an important aspect of writing and language development (see Schmitt, Chapter 15, this volume), plays a critical role "in developing use of morphosyntactic structures and text-production abilities" (p. 16) that goes beyond vocabulary complexity, density, and diversity in isolation, frequent measures for vocabulary development. In turn, "syntactic constructions … cannot be fully evaluated without taking account of lexical content" (p. 16), an orientation toward syntax that points to the need to consider them together in terms of lexicogrammar, rather than separately. Finally, "genre turns out to be an extremely relevant factor in relation to age differences in text production" (p. 31) and "mature speaker-writers deploy a wider variety of lexical, syntactic, and rhetorical devices for differentiating stance; and they also make more metatextual commentary" (p. 37).

While the Berman and Verhoeven study provided little information about the educational context from which the composition data were taken, exactly that focus drives studies hailing from the Australian K–12 environment that expressly intends to foster native language learning. Writing in different genres takes on a central role in a remarkable number of these studies, of which Christie (2012) is a particularly instructive, because comprehensive treatment. Though the few publications that I am here able to cite refer to genre families of different sizes and orientations (e.g., stories, histories, reports, explanations, procedures in Martin & Rose, 2005; the recording, explaining, and arguing genres as representative of the development of language development in the area of historical discourse throughout the school years in Coffin, 2006), they share an explicitly language-based notion of writing development understood as language development that requires the expansion of those language resources that are necessary for textual meaning-making in a variety of genre and disciplinary areas. Not only do they provide a plethora of suggestions for the kinds of macro-developmental issues that longitudinal study of writing and language learning can and should foster in educational settings (Schleppegrell, 2004); they do so by drawing on the rich conceptual and analytical apparatus provided by SFL, a textual-functional theory of language that is increasingly seen as uniquely beneficial for research practice at all educational levels, in highly diverse educational settings for a range of learners, and with different learning foci that, nevertheless, prominently include literacy development.

Finally, moving to foreign language instruction that has a strong longitudinal orientation, I refer to studies of L2 literacy development for EFL in the secondary school setting (see e.g., Whittaker & McCabe, Chapter 13, this volume; and Whittaker, Llinares, & McCabe, 2011), and the work done in my home department that comprises an entire four-year undergraduate program in German, expansively described in Byrnes, Maxim, and Norris (2010). I have highlighted the program's overall indebtedness to the fundamentally dynamic and developmentally oriented quality of SFL constructs as a way of investigating both advances in L2 writing and language learning (Byrnes, 2012, 2014, 2020a); and I have detailed development in both writing and language abilities from a textual-functional perspective. For example, I have longitudinally traced the gradual emergence of ideational meaning-making through the pivotal construct of 'grammatical metaphor' (Byrnes, 2009), especially when it is realized through nominalizations (see also Whittaker & McCabe, Chapter 13, this volume). In turn, the interpersonal metafunction in textual meaning-making is the focus of Ryshina-Pankova's (2011) study, which uncovered a trajectory from "foregrounding overt expression of the writer's opinion … towards a more intersubjective and indirect expression of stance that occurs through thematization of the putative reader" (p. 253) and suggests that "the ability to argue persuasively might be contingent both on the exposure and on engagement with academic argumentative genres and on the development of linguistic resources that enable one to take an intersubjective position" (p. 254).

To sum up, when longitudinal inquiry into the L2 writing-language learning interface focuses on the functional elaborateness of both writing and continued language learning, research should be able to uncover evidence for the intricately interwoven and potentially mutually constitutive nature of that connection that has, to this point, largely eluded us.

Privileging an extended curricular framework for researching the writing-language learning interface

With those reflections, the groundwork has been laid for the fourth perspective that might influence an agenda for research into the L2 writing – language learning interface. It reprises the educational and longitudinal imperatives of future research as I have presented them, but gives them additional focus through the well-established educational construct of curriculum. I am aware that for many a reader 'curriculum' is not a go-to construct that speaks of complexly conceptualized notions of emergent learning over the long time. In fact, most often the opposite is the case, inasmuch as curriculum is frequently seen as an administrative-structural convenience that assures static approaches to the delivery of instruction (see Manchón & Leow,

Chapter 14, this volume). In that interpretation it has little of the intellectual heft with which I have endowed it in a number of publications in order to acknowledge the significant impact on learning that assumptions about learning and development that animate a program have on that learning (see Byrnes, 2019). Implied in that assertion is that the explicit creation of such a curricular framework has the potential of making operationalizable the otherwise vague notion of 'context' of much research.

Readers may also find vexing the aspirational quality of my proposal to make curricular thinking a central component for a future research agenda that illuminates the writing – language learning interface. While that impression may be justified inasmuch as comprehensive curriculum construction for an entire program is rare; nevertheless, such a proposal is by no means unrealistic given that the curriculum to which I am referring by now draws on more than two decades of successful educative work for the benefit of students' learning and has been well documented in my own and my colleagues' publication record. For details I refer interested readers to ongoing work in curricular renewal <https://german.georgetown.edu/undergraduate/curriculum> and a listing of publications referencing that context <https://german.georgetown.edu/bibliography>. In particular, I have detailed its benefits for: setting up an integrated literacy-oriented program that explicitly uses writing to enhance language learning over the four-year duration of U.S. undergraduate studies <https://german.georgetown.edu/undergraduate/roleofwriting>; Byrnes, 2012, 2017); for details on its underlying assumptions, see Byrnes, 2020a, for its developmental trajectory, see <https://german.georgetown.edu/page/1242716542088.html>); for affirming the educational, as against the theoretical, motivation behind 'task' by linking it to genre (2002, 2014, 2015), both constructs that link writing and language learning; and, most broadly, by portraying curriculum as a way of specifying the instructed context of SLA (2019) in such a way that it can be researched with a strong developmental orientation that also espouses principles of CDST (Byrnes, 2020b).

Research that is tied to the particulars of an educational setting inevitably invites concerns about the extent to which its findings can apply to other contexts. But, given that we can no longer deny the highly contextualized and complexly dynamic nature of the relationship between L2 writing and learning, the more fundamental question is this: Can inquiry that lays out factors that provide reasoned accounts for the learning that can be observed within a particular educational setting be sufficiently insightful to be transferable to other contexts? Or, as Dörnyei, taking a dynamic systems perspective (2014, pp. 89–90) words the dilemma, are "the emerging prototypes and outcome patterns … sufficiently robust to be recognizable in a variety of contexts?" I believe not only that they are, but that they constitute a more honest, a more valid, and a more actionable way to deal with

instructed language use phenomena, particularly when research adopts an ecological perspective that recognizes the relational quality of all educational learning and accords to the particularities of a given context the power to shape that learning.

Concluding reflections

On that note, my concluding reflections reprise two areas that have resonated throughout this chapter but merit special mention.

My first point is that the desired future research agenda would be enormously facilitated by a functional theory of language, that is, a meaning-oriented, textual theory of language. Such a claim seems commonsensical not least because such a theory, in contrast with the 'hidden' structuralist theory of language that prevails in much writing research, will have had to give considerable thought to exactly what it is about language as a social-semiotic system of meaning-making that has made its formal system what it is. Presumably, constructs and processes that are laid out in that pursuit would be prime candidates for guiding future research. Readers will not be surprised if I reiterate a position that I have taken throughout this chapter and have extensively addressed in a number of publications that SFL presents a particularly advantageous choice. Beyond its unmistakable meaning and text orientation, it recommends itself due to (a) its expansive analysis of the entire system of various languages, as expected, particularly of English (Halliday/Matthiessen, 2014); (b) its focused attention on the relationships between cognition and language (Halliday, 2004; Halliday & Matthiessen, 1999); (c) its dynamic interpretation of the function of grammar in terms of complementarities that are open to user choices along the paradigmatic axis rather than as fixed rules that must be observed along the syntagmatic axis (Halliday, 2008); (d) its explicit attention to different modes of meaning-making in spoken and written language (Halliday, 2002); and, finally, (e) its extensive engagement with the impact of education on language development (Halliday, 1993, 1999) and ways of knowing (Byrnes, 2020b).

My second point is an understanding of social reality in terms of principles put forward in CDST (see, e.g., Larsen-Freeman, 2017). For the topic at hand, a particular challenge is that of reconciling the complexly dynamic nature of language use and development – its complex systems nature – with the unavoidably restricting exigencies of educational settings. Complexity theorists have addressed this concern in terms of the need for establishing boundaries, with Cilliers (2001) noting that "boundaries are simultaneously a function of the activity of the system itself, and a product of the strategy of description involved … The boundary of the system is therefore neither purely a function of our description, nor is it purely a natural thing" (p. 141). I consider the construct of curriculum to be one way of recognizing

these countervailing dynamics: On the one hand, the need to assure agency, choice, and creativity on the part of teachers, learners, and program administrators; on the other hand, the need for a conceptual frame of reference for how things will 'be done' in a program that acknowledges that education, like politics, will always have a certain normative (and power) dimension by way of the goals and approaches it privileges as desired 'end points'.

Reflections like these bring to mind the proposal by van Geert and Steenbeek (2014) who, from a complex systems perspective, allow for the possibility of considering curriculum as a praxis-based 'simplex system' within the overall epistemic complexity of the larger educational setting. It addresses the need on the part of educators to understand with an experientially-based, holistic sense of their educational setting how they might accomplish valued educational goals; their additional need to evaluate the educational processes they have privileged; and, finally, to shape them discursively for themselves and for their diverse publics so that they become knowledge. There is an undeniable need to be able to 'account' in the multiple meanings of that word for educative actions and narratively to give them meaning and value while continuing to uphold the epistemic complexity of the system as expressed by the curricular framework (Byrnes, 2019, 2020b).

Because it is impossible to know the entire complex system, researchers will need to work collaboratively with educators in order to learn what it means to learn how to write in and learn additional languages in complexly conceptualized educational settings. From that experience they might draw both humility regarding their claims as well as a deep sense of being able to make worthwhile contributions to some of the most pressing concerns in contemporary societies. No longer stifled by the untenable demands of an objective rationality, they might find themselves liberated to engage in a self-critical rationality that foregrounds imagination and action toward a more flourishing social future (Preiser & Cilliers, 2010). To contribute to that possibility for the L2 writing research agenda of the future has been among the goals of this chapter.

References

Berman, R. A., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language and Literacy*, 5, 1–43. https://doi.org/10.1075/wll.5.1.02ber

Biesta, G. (2010). Five theses on complexity reduction and its politics. In D. Osberg & G. Biesta (Eds.), *Complexity theory and the politics of education* (pp. 5–13). Rotterdam: Sense Publishers. https://doi.org/10.1163/9789460912405_003

Byrne, D., & Callaghan, G. (2014). *Complexity theory and the social sciences. The state of the art*. New York, NY: Routledge.

Byrnes, H. (2002). The role of task and task-based assessment in a content-oriented collegiate FL curriculum. *Language Testing*, 19, 419–437. https://doi.org/10.1191/0265532202lt238oa

Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education*, 20, 50–66. https://doi.org/10.1016/j.linged.2009.01.005

Byrnes, H. (2012). Conceptualizing FL writing development in collegiate settings: A genre-based systemic functional linguistic approach. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 190–218). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781934078303.191

Byrnes, H. (2014). Linking task and writing for language development: Evidence from a genre-based curricular approach. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 237–263). Amsterdam: John Benjamins.

Byrnes, H. (2015). Linking 'task' and curricular thinking: An affirmation of the TBLT educational agenda. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 193–224). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.8.07byr

Byrnes, H. (2017). Meeting the challenge of instructed language development. Reflections on systemic-functional contributions. In S. Neumann, R. Wegener, J. Fest, P. Niemietz, & N. Hützen (Eds.), *Challenging boundaries in linguistics: Systemic functional perspectives* (pp. 457–491). Frankfurt: Peter Lang.

Byrnes, H. (2019). Affirming the context of instructed SLA: The potential of curricular thinking. *Language Teaching Research*, 23, 176–183. https://doi.org/10.1177/1362168818776666

Byrnes, H. (2020a). Developing an SFL-inspired 4-year collegiate foreign language program: Contexts, constructs, curriculum. In J. R. Martin, J. S. Knox, & D. Caldwell (Eds.), *Developing theory: A handbook in appliable linguistics and semiotics*. London: Bloomsbury.

Byrnes, H. (2020b). Envisioning writing development in CDST under a curricular optic: A proposal. In G. G. Fogal & M. Verspoor (Eds.), *Complex Dynamic Systems Theory and L2 writing development* (pp. 243–272). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.54.10byr

Byrnes, H., & Manchón, R. M. (2014a). Task, task performance, and writing development: Advancing the constructs and the research agenda. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 267–299). Amsterdam: John Benjamins.

Byrnes, H., & Manchón, R. M. (2014b). Task-based language learning – Insights from and for L2 writing: An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based L2 language learning – Insights from and for L2 writing* (pp. 1–23). Amsterdam: John Benjamins.

Byrnes, H., Maxim, H. H., & Norris, J. M. (2010). Realizing advanced foreign language writing development in collegiate education: Curricular design, pedagogy, assessment. *Modern Language Journal*, 94(Supplement s-1), iv–vi, 1–235.

Christie, F. (2012). *Language education throughout the school years: A functional perspective. Language Learning*, 62(Supplement 1).

Cilliers, P. (2001). Boundaries, hierarchies, and networks in complex systems. *International Journal of Innovation Management*, 5, 135–147. https://doi.org/10.1142/S1363919601000312

Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. London: Continuum.

Craik, F. I. M. (2002). Levels of processing: past, present, and ... future? *Memory*, 10, 305–318. https://doi.org/10.1080/09658210244000135

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511. https://doi.org/10.1177/0741088390007004003

Cumming, A. (2016). Theoretical orientations in L2 writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 65–88). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-006

Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modeling' in the language classroom. *Language Teaching*, 47, 80–91. https://doi.org/10.1017/S0261444811000516

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, 28(2), 122–128. https://doi.org/10.2307/356095

Halliday, M. A. K. (1973). *Explorations in the functions of language*. London: Edward Arnold.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education*, 5, 93–116. https://doi.org/10.1016/0898-5898(93)90026-7

Halliday, M. A. K. (1999). The notion of "context" in language education. In M. Ghadessy (Ed.), *Text and context in functional linguistics* (pp. 1–24). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.169.04hal

Halliday, M. A. K. (2002). Spoken and written modes of meaning. In J. J. Webster (Ed.), *On grammar* (Vol. 1, pp. 323–351). London: Continuum.

Halliday, M. A. K. (2004). Things and relations: Regrammaticising experience as technical knowledge (1998). In J. J. Webster (Ed.), *The language of science* (Vol. 5, pp. 49–101). London: Continuum.

Halliday, M. A. K. (2008). *Complementarities in language*. Beijing: The Commercial Press.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999). *Construing experience through meaning: A language-based approach to cognition*. London: Continuum.

Halliday, M. A. K., revised by Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar*. New York, NY: Routledge. https://doi.org/10.4324/9780203783771

Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers. An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly*, 8, 229–249. https://doi.org/10.1080/15434303.2011.565844

Larsen-Freeman, D. (2011). A complexity theory approach to second language development/acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 48–72). New York, NY: Routledge.

Larsen-Freeman, D. (2017). Complexity theory: The lessons continue. In L. Ortega & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 11–50). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.48.02lar

Leow, R. P. (2019). From ISLA > ISLA > ILL: A curricular perspective. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 485–493). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. P., & Cerezo, L. (2016). Deconstructing the "I" and "SLA" in ISLA: One curricular approach. *Studies in Second Language Learning and Teaching*, 6, 43–63. https://doi.org/10.14746/ssllt.2016.6.1.3

López Serrano, S., Roca de Larios, J., & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically motivated and empirically-based coding system. *Studies in Second Language Acquisition*, 41(3), 503–527. https://doi.org/10.1017/S0272263119000275

Macaro, E. (2014). Reframing task performance: The relationship between tasks, strategic behaviour, and linguistic knowledge in writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 53–77). Amsterdam: John Benjamins.

Manchón, R. M. (2009). Introduction: Broadening the perspective of L2 writing scholarship: The contribution of research on foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 1–19). Clevedon, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859-004

Manchón, R. M. (2011). Writing to learn the language: Issues in theory and research. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 61–82). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.07man

Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing processes. Implications for theory. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching and research* (pp. 102–124). Clevedon, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859-008

Martin, J. R., & Rose, D. (2005). *Genre relations*. London: Equinox.

Norris, J. M., & Manchón, R. M. (2012). Investigating L2 writing development from multiple perspectives: Issues in theory and research. In R. M. Manchón (Ed.), *Writing development: Multiple perspectives* (pp. 219–245). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781934078303.221

Osberg, D. (2008). The logic of emergence: An alternative conceptual space for theorizing critical education. *Journal of the Canadian Association for Curriculum Studies*, 6, 133–161.

Preiser, R., & Cilliers, P. (2010). Unpacking the ethics of complexity: Concluding reflections. In P. Cilliers & R. Preiser (Eds.) *Complexity, difference and identity: An ethical perspective* (pp. 265–287). Dordrecht: Springer. https://doi.org/10.1007/978-90-481-9187-1_13

Roca de Larios, J. (2013). Second language writing as a psycholinguistic locus for L2 production and learning. *Journal of Second Language Writing*, 22, 444–445. https://doi.org/10.1016/j.jslw.2013.08.009

Roca de Larios, J., Nicolás-Conesa, F., & Coyle, Y. (2016). Focus on writers and processes. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language learning* (pp. 267–286). Berlin: de Gruyter Mouton. https://doi.org/10.1515/9781614511335-015

Ryshina-Pankova, M. (2011). Developmental changes in the use of interactional resources. Persuading the reader in FL book reviews. *Journal of Second Language Writing*, 20, 243–256. https://doi.org/10.1016/j.jslw.2011.05.011

Ryshina-Pankova, M., & Byrnes, H. (2013). Writing as learning to know: Tracing knowledge construction in L2 German compositions. *Journal of Second Language Writing*, 22, 179–197. https://doi.org/10.1016/j.jslw.2013.03.009

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410610317

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate. A step towards second language learning. *Applied Linguistics*, 16, 371–391. https://doi.org/10.1093/applin/16.3.371

van Geert, P., & Steenbeek, H. (2014). The good, the bad and the ugly? The dynamic interplay between educational practice, policy and research. *Complicity: An International Journal of Complexity and Education*, 11(2), 22–39. https://doi.org/10.29173/cmplct22962

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: The MIT Press.

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, 15, 343–362. https://doi.org/10.1177/1362168811401154

Witte, S. P. (1987). Pre-text and composing. *College Composition and Communication*, 38, 397–425. https://doi.org/10.2307/357634

L2 writing-to-learn

Theory, research, and a curricular approach

Ronald P. Leow

Georgetown University

Given the current contextual focus on the cognitive processes employed during L2 written production in ISLA with a focus on L2 *writing-to-learn* in relation to written corrective feedback (WCF), this chapter provides (1) a description of the curricular context in which the *writing-to-learn* perspective lies, together with its associated learning outcomes and assumed pedagogical ramifications derived from empirical research; (2) a succinct review of theoretical underpinnings postulated to account for cognitive processes employed during the phase of addressing WCF; (3) a critical report of L2 WCF studies premised on these underpinnings, their curricular association, and potential pedagogical extrapolations; and (4) a curricular approach that involves potential directions for future research premised on understanding and promoting such processes during this writing process in relation to L2 development.

Introduction

Writing is a visual form of communication that comprises an important component in many language curricula (see, for example, short or long written assignments to be performed at home or in a testing condition). It is considered a complex *productive skill* given that to share information, writers need to successfully manipulate their current knowledge of the second or foreign language (L2) in order to create a text with new or previously learned language for a potential reader. At the same time, many adult L2 writers appear to face many challenges, most likely due to many variables, in the entire L2 writing process.

There have been several theoretical underpinnings postulated to account for the L2 writing process (Manchón, 2012) and empirical studies have addressed the linguistic processing involved in composing the L2 (e.g., Cumming, 1990; Manchón, Roca de Larios, & Murphy, 2009; Roca de Larios, Murphy, Manchón, & Marín, 2008; López-Serrano, Roca de Larios, & Manchón, 2019). Indeed, the L2 writing

process has been of tremendous interest, importance, and debate for decades in the instructed second language acquisition (ISLA) strand of research (see Manchón & Vasylets, 2019; Polio & Lee, 2017; Roca de Larios, Coyle, & Nicolás-Conesa, 2016; Wang & Jiang, 2015, for recent reviews). However, there has been a subtle shift in the perspective of the writing process to probe deeper into the potential for learners' overall L2 development being derived from this process. This shift has resulted in a distinction between *learning-to-write* (mostly associated with writing courses) and *writing-to-learn* (mostly situated within an L2 curriculum) and between "feedback for accuracy" (effect on immediate performance) versus "feedback for acquisition" (changes in the L2 system) (Manchón, 2011a).

The phrase *writing -to-learn* has been around for quite a while and has been defined, for example, as "a strategy through which students can develop their ideas, their critical thinking ability and their writing skills. Writing to learn enables students to experiment every day with written language and increase their fluency and mastery of written conventions" (Texas Computer Education Association (TCEA) 2013: Strategies for Engaging the 21st Century Learner). *Writing-to-learn* activities have been described as "short, informal writing tasks that help students think through key concepts or ideas central to a course. Quite often, these activities require very little class time or can be assigned as short, out-of-class writing their ideas" (The Department of English, Modern Foreign Languages and Speech Communication, Jackson State University). While these definitions appear to be situated within a *learning-to-write* perspective, early ISLA studies in the 90s (e.g., Cumming, 1990; Swain & Lapkin, 1995), following the role of output postulated by Swain (1985), focused more on a *writing-to-learn* perspective that addresses specifically the potential of writing for L2 development during the initial stage of composing. For example, Swain and Lapkin (1995) observed, via think aloud protocols, that some of their participants were noticing and solving their language problems while composing. Cumming (1990) observed that writing can promote learner attention to form-meaning connections leading to a more solid knowledge of the L2. In addition, it was suggested that students interacting with the feedback provided by teachers might promote noticing and subsequent revision of the incorrect L2. This *writing-to-learn* perspective, raised over a decade and a half ago (e.g., Harklau, 2002), gained traction after the publication of two influential edited books (Manchón, 2009, 2011b) that underscored the potential of writing as an integral part of overall language instruction (Manchón, 2011c) and the elevation of the role of writing beyond mere practice of grammar and vocabulary and more as an opportunity to learn the L2 (Ortega, 2011). Recently, Manchón and Vasylets (2019), following Manchón and Williams (2016), underscored the need for "research interest in ascertaining how and why such rich linguistic processing may be beneficial

in terms of language development" (p. 342), directly relating the processes involved in composing with overall L2 development (see also Introduction to this volume).

Writing-to-learn associated with potential L2 learning is refocusing the role of written corrective feedback (WCF) in current ISLA research, a refocus that logically lends some ecological validity to research attempting to address the role of WCF in the instructed setting. Theoretically, several cognitive underpinnings have been cited to account for the role or lack thereof of WCF in the L2 learning process. Empirically and methodologically, Manchón (2012) raised a central question that is clearly designed to push this strand of WCF research to probe deeper into the issue of "How can we observe L2 writing development and the factors that determine it?" (p. 222). Since then, there has been an increasing call not only to study the role of writing processes in direct relation to L2 development (e.g., Manchón & Vasylets, 2019; Manchón & Williams, 2016) but also a recent uptick in empirical studies methodologically attempting to address the cognitive processes assumed to take place while L2 learners compose and produce their thoughts in writing in relation to WCF (e.g., Adrada-Rafael & Filgueras-Gómez, 2019; Caras, 2019; and Park & Kim, 2019 for online verbal reports, and Cerezo, Manchón, & Nicolás-Conesa, 2019; Manchón, Nicolás-Conesa, Cerezo, & Criado, 2020, and Suzuki, 2012, 2017 for offline written languaging; see also Suzuki & Storch, 2020).

From a curricular perspective, it has been suggested that the writing process should be situated within the context in which it occurs, namely, the instructed setting, that logically includes both L2 learners and instructors (Manchón, 2012; Manchón & Vasylets, 2019) and, as an extension, the expected learning outcomes of any given L2 curriculum (see Leow, 2019a, 2019b; Leow & Cerezo, 2016). Situating the writing process within such a curricular approach logically places a premium on the value of pedagogical implications derived from empirical ISLA research, based on its curricular status (Leow & Cerezo, 2016).

Given the current contextual focus on the cognitive processes employed during L2 written production in ISLA with a focus on L2 *writing-to-learn* in relation to WCF, this chapter provides (1) a description of the curricular context in which the *writing-to-learn* perspective lies, together with its associated learning outcomes and assumed pedagogical ramifications derived from empirical research; (2) a succinct review of theoretical underpinnings postulated to account for cognitive processes employed during the phase of addressing WCF; (3) a critical report of L2 WCF studies premised on these underpinnings, their curricular association, and potential pedagogical extrapolations; and (4) a curricular approach that involves potential directions for future research premised on understanding and promoting such processes during this writing process in relation to L2 development.

Writing in the L2 curriculum

Understanding or minimally acknowledging the curricular role of L2 writing and its relationship to learning in academic institutions is of paramount importance if empirical research is slated to address this instructed environment, potentially contribute to a better understanding of the writing process (whether *learning-to-write* or *writing-to-learn*), and provide robust pedagogical ramifications to this formal setting. Underlying this knowledge is the need for a concise definition of what comprises learning and WCF.

First of all, there is a need to differentiate between *writing-to-learn* in a language curriculum and *writing-to-learn* in a writing course that is usually viewed as *learning-to-write*. A global overview of a typical language or writing curriculum and the roles written compositions play raises much caution in conflating these two levels of writing. For example, the objective of writing in many language classes is primarily to provide students with the opportunity to *practice* what has been covered in a chapter, lesson, or unit specifically in relation to grammatical point(s), vocabulary, and content (usually the topic of the chapter just covered). Crucially, writing is only one of the usual four skills being promoted in the language classroom and, in many language curricula, it is integrated with the other three skills, once again in relation to grammar, vocabulary, and content. Consequently, WCF may not be the only source of negative evidence or error correction in an instructed setting.

The length of the composition varies across levels of language proficiency, from, for example, half a page in a Beginning L2 class to over three pages in the writing course. The amount of time allocated to correcting compositions is clearly differential. Teachers in language classes, for example, at US college levels, typically grade compositions of varying lengths approximately every three weeks while teachers in writing courses grade compositions more often and with longer lengths. Feedback provided on compositions in both language and writing courses may be componential, that is, several categories (e.g., writing mechanics, structure, content, vocabulary, grammar) are addressed, or a combination of these categories with grammar being the most popular or weighted category. Writing courses also provide feedback on successful completion of different types of genre. In other words, while many WCF studies appear to focus almost exclusively on grammatical or linguistic accuracy, reduction of grammatical errors, or "writing" or "writing ability", this may not be reflective of many language or writing courses. Indeed, WCF studies that only focus on one or two linguistic items in the composition appear to replicate grammar exercises (van Beuningen, 2010; Xu, 2009), missing the other writing features obviously involved in composing, including students' expectations.

Second, L2 curricula also have objectives for each skill and associated learning outcomes, an issue that needs to be seriously considered if one were to situate any WCF study within a language curriculum, especially at the college level. In other words, a simple focus on grammatical accuracy does not fall within the objectives and learning outcomes of a typical language curriculum. From a learning perspective, it is also important to clearly operationalize what comprises learning and address whether such "learning" reported in WCF studies was robust enough to warrant pedagogical extrapolation to the L2 classroom.

Third, it is important to define what WCF is and its intended purpose. There are several definitions (e.g., Bitchener & Storch, 2016; Ellis, 2009) but WCF can be viewed as any external manipulation of L2 writers' product by the teacher or the researcher designed to minimally draw their attention to some grammatical, lexical, structural, and/or content error committed by the L2 writers.[1] From a grammatical or lexical perspective, WCF may take the form of (a) any number of type of feedback that includes direct (providing the exact correction), indirect (underlying the error, reformulations, models), and metalinguistic (providing a coding of type of error), or (b) amount of feedback (focused vs. unfocused). More importantly, WCF is clearly premised on explicit learning grounded in the activation of prior explicit knowledge (Leow, 2015; Polio, 2012) and the hope that L2 writers not only pay attention to or notice WCF but also adequately further process (with some depth of processing or level of awareness) the information provided in the feedback with the goal of restructuring their incorrect L2 knowledge. Within this explicit learning condition, any argument against the provision of WCF in relation to the development of implicit knowledge (e.g., Truscott, 1996), reference to type of knowledge (explicit vs. implicit) employed (e.g., Bitchener & Knoch, 2010; Shintani & Ellis, 2013), or the process of acquisition[2] may be moot. Given that WCF is external and the need to establish whether the L2 writer adequately processes the feedback provided or even understands it, it is of paramount importance that learner processes during this revision phase of the writing process are investigated fully in an effort to understand better how WCF plays a role in subsequent restructuring and potential learning. In turn, these concurrent data can be used to provide the link between the writing processes and potential L2 development.

---

1.   It may be argued that both textual enhancement (see Leow & Martín, 2017 for a recent review) and WCF share this premise with the difference lying in the former being receptive (textual enhancement) and the latter both receptive and productive.

2.   See Leow (2015, 2019a) for a distinction between learning and acquisition based on contextual (naturalistic vs. classroom), type of classroom processing (implicit vs. explicit learning), and curricular dimensions.

The next section provides succinct reports of theoretical underpinnings postulated to account for how L2 writers process WCF with the aim of underscoring the important role of cognitive processes and how these processes are accounted for in each underpinning.

Theoretical underpinnings for WCF

As discussed above, WCF, irrespective of type or amount, is provided within an explicit learning condition premised on conscious processing, processes, and primarily explicit knowledge. Subsequently, only four theoretical underpinnings cited in ISLA may account for explicit learning during the revision phase of the writing process, namely, Schmidt's (1990) Noticing Hypothesis, Swain's (2005) Output Hypothesis, DeKeyser's (2015) Skill Acquisition Theory, and Leow's (2015) Model of the L2 Learning Process in ISLA.[3] Each is described briefly below together with their contribution to a better understanding of how or why WCF can lead to language learning.

Schmidt's Noticing Hypothesis

Schmidt's (1990 and elsewhere) Noticing Hypothesis was originally situated at the early stage of the L2 learning process (the input-to-intake stage) with an apparent focus on new linguistic information in the L2 input. Noticing is defined as attention accompanied by minimally a low level of awareness (focal attention being isomorphic with awareness) and it is also posited that learning without awareness, while somewhat possible, has minimal chances of success.

In addition to noticing, Schmidt distinguishes a higher level of awareness, which he calls "understanding" and which is related to the ability to analyze, compare, and test hypotheses about the linguistic input leading to rule formulation. Schmidt postulates that while noticing is necessary for intake and potential learning to take place, awareness at the level of understanding may act as a facilitator for learning, but its presence is not necessary. The crucial difference between noticing and understanding is that the former results in intake and in item learning while the latter leads to restructuring and system learning.

---

3.   Several researchers (e.g., Manchón & Vasylets, 2019; Van Beuningen, 2010; Wang & Jiang, 2015) have included *focus-on-form* (Long & Robinson, 1998) as a theoretical underpinning to account for the role of WCF in L2 development. However, the principles of *focus-on-form* that rest on some periodic focus on grammatical errors during communication do not adhere easily to the written mode.

Schmidt's Noticing Hypothesis underlies researchers' assumption that L2 writers need to minimally "notice" feedback before any potential and subsequent restructuring of their interlanguage may take place. Of interest is that employing Schmidt's Noticing Hypothesis at the output stage of the L2 learning process and not at the original initial stage of input processing to account for the role of feedback fails to address several other potential variables that may play a role at this later stage. However, at this output stage, noticing has to be necessarily linked to the L2 writer's prior knowledge. In other words, the L2 writer notices the mismatch between what she has produced (from prior knowledge already lodged in her internal system) and the different linguistic information contained in the WCF. However, to process the information contained within the WCF, the L2 writer needs to be more cognitively engaged, that is, beyond mere noticing, in order to potentially restructure her prior incorrect knowledge. This is not awareness at the level of noticing that is aligned with a relatively low depth of processing (Leow, 2012, 2015), but a higher level of awareness that involves deeper processing. Indeed, noticing (attention plus a low level of awareness) does not guarantee automatic further processing, deeper processing, or higher level of awareness. Both concurrent think aloud (e.g., Leow, 2001) and eye-tracking (e.g., Godfroid, Boers, & Housen, 2013) procedures have revealed the failure of noticing to result in subsequent performance (recognition and production) after the experimental exposure, indicating that intake may need to be further processed for potential learning to take place (Leow, 2015). Consequently, the Noticing Hypothesis may be too coarse-grained to account for how the WCF is processed after noticing and what role prior knowledge plays in any interaction with the WCF noticed.

Swain's Output Hypothesis

Swain's (2005) Output Hypothesis is situated at the late productive stage of the L2 learning process and was originally rooted in oral production but currently applied also to the written mode. The hypothesis makes three major claims regarding the functions of learner production during the learning process. The first two claims are psycholinguistic in nature while the third is sociocultural. The first claim is associated with the actual process of producing the L2 and potential cognitive processes involved during this production while the second, discussed in this chapter, is associated with cognitive processes associated with the potential role of feedback.

Swain's second claim, associated with a hypothesis-testing function, is related to learners interested in experimenting with new forms and structures and verifying whether they are correct or need correction. According to Swain, learners need to test hypotheses, which in turn could be confirmed or disconfirmed following

feedback. This process allows learning to take place given that consciously processing the feedback can potentially lead to modifying or "reprocessing" their output. Note, however, that the notion of "reprocessing" the feedback arguably weakens the primary tenet of the Output Hypothesis given that it will need to acknowledge the role of input processing in the L2 learning process.

Swain's Output Hypothesis is unique in the sense that it is situated at the output stage of the L2 learning process, addresses output or production (considered a product) yet claims that "the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning" (Swain, 2005, p. 471). In other words, Swain includes this process of producing as part of the learning process, during which quite a high depth of processing appears to be involved in all her three claims. The Output Hypothesis lends itself quite easily to the role of WCF given that it is situated primarily at the output stage of the L2 learning process, which is where WCF is provided. According to the Hypothesis, it is the hypothesis-testing function that affords the benefits of WCF although the recipients may be restricted to only those L2 learners who are highly motivated to process deeply any feedback provided on their compositions in relation to addressing their original hypotheses on the L2. One potential drawback is the delayed provision of WCF, which may impact L2 writers' hypotheses made during the composing phase. Also, how the feedback is further processed and whether L2 learning rests solely on production are questions that remain unanswered in this hypothesis.

Skill Acquisition Theory

DeKeyser's (2015) Skill Acquisition Theory views learning as a new skill that undergoes three developmental stages: declarative, procedural, and automatic. The theory focuses on how declarative knowledge (assumed to be accurate) carries learners through the proceduralization stage by way of carefully formatted tasks, to more implicit or procedural knowledge and into the initial stages of automatization. Different types of knowledge may be achieved differentially although, under this theory, a learner cannot reach a practical proficiency level without moving through each stage.

The Skill Acquisition Theory, like the Output Hypothesis, lies at the output stage of the L2 learning process and is primarily associated with accurate explicit prior knowledge and how to alter such knowledge to a less explicit one, namely, procedural. It may fall within the belief that WCF provides learners the opportunity to practice the L2 more accurately until they convert declarative knowledge into procedural knowledge (e.g., Bitchener, 2016). This belief, however, does not truly reflect the type or degree of practice provided in the writing component of

a language curriculum. Leeman (2007) provides some suggestions regarding the potential role feedback may play at the three stages postulated within the theory. At the initial stage, feedback can promote the development of declarative knowledge while during the stages of proceduralization and automatization, feedback can "indicate the need for greater attention and reliance on declarative knowledge as well as the need to change the scope of a given rule or procedure. Furthermore, feedback may be useful in avoiding the automatization of non-target L2 knowledge" (Leeman, 2007, p. 117). While it may be proposed that feedback associated with subsequent corrected and consistent practice promotes declarative L2 knowledge during the three stages, there is an assumption that the declarative knowledge being practiced is somewhat inaccurate and needs to be modified via WCF during practice. At the same time, note that Skill Acquisition Theory is premised on real and constant practice using declarative knowledge (assumed to be accurate) during the proceduralization stage until it is automatized or proceduralized. Consequently, whether WCF plays an important role in this theory is not well explicated nor is it well explained how L2 writers process WCF.

The model of the L2 learning process in ISLA

Leow's (2015) model views the notion of learning as consisting of both processes and products and elaborates on three processing stages (input processing, intake processing, knowledge processing) in light of the cognitive processes postulated to play important roles during these stages. This model is crucially situated within the instructed setting and while attention is central to the model, depth of processing, defined as "the relative amount of cognitive effort, level of analysis, elaboration of intake together with the usage of prior knowledge, hypothesis testing and rule formation employed in decoding and encoding some grammatical or lexical item in the input" (p. 204) plays an important role at several stages of the L2 learning process. Of the three processing stages, the knowledge processing stage is most pertinent to WCF, as described below.

The knowledge processing stage is the third and final processing stage that occurs between the L2 developing system and what is produced by the learner (*knowledge processing*, e.g., assigning phonological features to the L2 in oral production, monitoring production in relation to learned grammar etc.). Depth of processing and potential level of awareness may play a role at this stage together with the ability to activate (appropriate) knowledge. At this stage, the model provides a loop for feedback to be further processed by the L2 learner as new input.

Like Swain's (2005) Output Hypothesis, Leow's model views the learning process to include the knowledge processing stage that allows potential feedback to

loop back to the early input processing stage. According to the model, whether the feedback has been attended to, detected, or noticed once again depends upon the attentional resources allocated to the feedback by the learner in addition to the depth of processing and level of awareness involved to make the connection between the learner's prior inaccurate knowledge or output and the information in the feedback received. In other words, whether the feedback processing allows for potential restructuring of the inaccurate knowledge may depend upon how deeply the feedback is processed or the level of awareness in relation to the mismatch between the learner's prior knowledge and the feedback.

Based on his Model of the L2 learning process in ISLA, and premised on his postulations on the different stages of the L2 learning process, the following fine-grained feedback processing framework is proposed, as seen below in Figure 1:
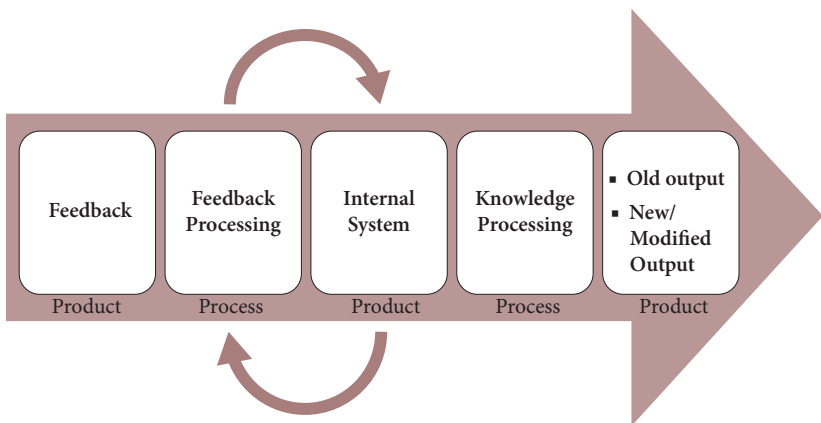

Figure 1. Feedback processing framework based on Leow's (2015) Model of the L2 Learning Proscess in ISLA

Feedback on learners' output is the L2 information that learners need to minimally pay attention to for feedback intake (attended, detected, or noticed) to enter into the learner's working memory. *Feedback Processing* encompasses *how* the learner cognitively processes the feedback (if at all) in relation to the current learner knowledge or interlanguage. If further processed at this stage, whether with a low or high depth of processing or level of awareness, the information in the feedback allows for reinforcement of accurate prior knowledge or, based on corrective feedback, for the potential of restructuring of previously learned inaccurate knowledge stored in the learner's *Internal System*. The new restructured information (accurate or still inaccurate) then replaces or joins the original knowledge in the *Internal System*, which is then available for the *Knowledge Processing* stage. There is, then, the possibility that the learner still retains the previous inaccurate L2 data and now holds

both (accurate and inaccurate) options in the system. Old (or inaccurate) output represents a potential absence or low depth of prior processing of the corrective feedback provided or not much confidence in the newly restructured knowledge if the feedback was indeed internalized. *New* or *Modified Output* is the learner's production of the restructured L2 and assumed to represent the L2 knowledge (as a chunk of language/item learning or systemized) the learner has at that point in time in her internal system. Delayed performance may indicate whether a complete accurate restructuring took place (as in system learning) or whether such restructuring was temporary or immediate or reflective of item learning, that is, accurate performance was evidenced immediately after the feedback was provided but over time the learner reverted back to her previous inaccurate interlanguage. Whether feedback is indeed processed by L2 learners may depend on several cognitive processes and variables that include depth of processing, levels of awareness, activation of appropriate prior knowledge, hypothesis testing, rule formulation, and/or metacognition. This framework provides a cognitive explanation for the role of corrective feedback, whether oral or written, in L2 development in direct relation to how L2 learners or writers process such feedback.

The next section provides a concise report on relevant empirical studies on WCF and L2 development. As will be seen, the bulk of this strand of research focused primarily on the product derived from WCF, failing to provide much needed insights into *how* L2 writers processed the feedback provided.

Empirical studies

The potential for L2 development to be derived from the writing process, especially in relation to the role of WCF for errors produced on students' compositions, has been empirically investigated and debated for decades (see Bitchener & Storch, 2016; Manchón & Vasylets, 2019, for recent reviews). What is of concern in the debate over its efficacy in L2 development is the conflation of the two environments, namely, language curriculum versus writing curriculum, as evident in the debates between Truscott (1999, 2004) versus Ferris (1999) and Chandler (2003), respectively.

While many studies report conflicting findings, it is instructive to review specifically whether these studies addressed *how* L2 writers processed the feedback provided, what L2 writers "learned" from these interventions, and whether the research design was linked to the language curriculum. To keep the focus of *writing-to-learn* within a language curriculum, studies employing populations from writing courses are not included in this review.

Direct and indirect feedback

Studies comparing types of WCF varying in directness have produced conflicting results. Some studies have reported support for direct WCF (e.g., Chandler, 2003, study 2; López, Van Steendam, Speelman, & Buyse, 2018; van Beuningen, De Jong, & Kuiken, 2008), no support (e.g., Semke, 1984; Shintani & Ellis, 2013), and no difference between the two (e.g., Ferris, 2006; Robb et al., 1986). At this point, it is not possible to make any claims about the efficacy of feedback directness on L2 learners' written accuracy. These inconsistencies could be rooted in the incomparability of research designs, methodological shortcomings and/or the possible interaction of different factors as confounding variables (i.e., classroom instruction or outside exposure) due to the quasi-experimental nature of the research design, task type, type of linguistic item, the participants' proficiency level, or the duration of the experimental exposure. Several intervening variables have also been identified that may impact the effectiveness of WCF, such as type of targeted error (e.g., Shintani, Ellis, & Suzuki, 2014), L2 proficiency or meta-linguistic awareness (e.g., Hyland & Hyland, 2006), or type of knowledge (partial vs. new) (Ellis, Sheen, Murakami, & Takashima, 2008).

Several researchers (e.g., Ashwell, 2000; Bitchener & Knoch, 2008; Ferris & Roberts, 2001; Lalande, 1982) have assumed that indirect WCF benefits L2 writers more given that they need to be cognitively engaged while processing the feedback: "it requires pupils to engage in guided learning and problem solving and, as a result, promotes the type of reflection that is more likely to foster long-term acquisition" (Bitchener & Knoch, 2008, p. 415). On the other hand, other researchers (e.g., Bitchener & Ferris, 2012; Chandler, 2003; Ellis et al., 2008; Sheen, 2007; Van Beuningen, 2010) view the absence of direct feedback as detrimental to the learning process due to the lack of linguistic feedback that can confirm or disconfirm their original hypotheses on the L2, especially complex errors such as syntax. In their opinion, direct WCF does provide the explicit grammatical information immediately that allows L2 writers to test their hypotheses about the L2 and internalize the correct form during the revision phase.

Unfocused and focused feedback

Unfocused feedback studies (e.g., Ashwell, 2000; Cerezo, Manchón, & Nicolás-Conesa, 2019; Chandler, 2003, study 2; Fathman & Whalley, 1990; Ferris, Liu, Sinha, & Senna, 2013; Kepner, 1991; Lalande, 1982; Robb et al., 1986; Semke, 1984) or "comprehensive" feedback in some studies (e.g., van Beuningen, De Jong, & Kuiken, 2012) have addressed the effect of unfocused feedback within a permutation of type

of feedback (direct, indirect, and metalinguistic) and provided overall inconsistent results on the effectiveness of amount of WCF.

In contrast, focused feedback studies (e.g., Benson & DeKeyser, 2019; Bitchener, 2008; Bitchener & Knoch, 2008, 2010; Ellis et al., 2008; Karimi, 2016; Sheen, 2007; Shepherd, Daily O'Meara, & Snyder, 2016; Shintani et al., 2014; Stefanou & Revesz, 2015), also conducted within a permutation of type of feedback, have offered overall a more positive result not only for performances on the immediate posttest but also on delayed posttests, with the caveat that several of these studies investigated only English articles. This relatively narrow focus on English articles gave rise to several studies (e.g., Benson & DeKeyser, 2019; Kassim & Ng, 2014; Shintani et al., 2014; Van Beuningen et al., 2012) that attempted to address the effect of WCF on type of linguistic items that include both rule-governed or treatable (e.g., verbs) or non rule-governed or untreatable (e.g., prepositions) or type of past tense form (present perfect vs. simple past). Overall findings appear to indicate a relatively complex situation that involves a permutation of several variables pertaining to the type of error (e.g., salience, complexity, error treatability) and type of feedback (direct, indirect, or metalinguistic) playing potential roles in whether or not WCF is successful.

Like the direct versus indirect comparison, different assumptions have been made with respect to the differential effectiveness of focused and unfocused WCF. For example, Ellis et al. (2008) assume that focused WCF when compared to unfocused WCF has greater potential not only to attract L2 writers' noticing (à la Schmidt, 1990) of the corrections but also lead to understanding them due to the limited amount of focused feedback provided. Sheen (2007) and Bitchener (2008) claim that unfocused WCF may not be the most effective correction method if L2 learners' limited processing capacity were considered. To this end, they assume that asking learners to deal with unfocused WCF that requires the simultaneous correction of a large range of linguistic features might lead to a cognitive overload while prohibiting learners' processing of the feedback. On the other hand, other researchers have queried the usefulness of focused WCF from both an ecological and writing process perspective. Storch (2010) noted that providing focused WCF goes against what is expected in the L2 classroom. Teachers do not target selected errors in students' compositions, which in turn may lead to some confusion on the students' part. Bruton (2009) raised the issue of whether such focused WCF might be viewed as more focus on form than belonging to the writing process.

In sum, the overall findings of WCF studies appear to indicate a clear advantage for feedback over no feedback. However, a critical review of WCF studies reveals major limitations in the research designs. First of all, with the exception of Caras (2019), Cerezo et al. (2019), Manchón et al. (2020), Coyle, Cánovas-Guirao, & Roca de Larios (2018), and Suzuki (2012, 2017), WCF studies have failed to employ

process measures to establish that participants were indeed not only paying attention to but also processing the feedback provided on their errors. Given that WCF is premised minimally on L2 writers' attention to the WCF, the failure to establish methodologically participants' processing of such WCF lowers the internal validity of these studies (Leow, 1999, 2015).

Second, it is quite challenging to establish what specifically participants learned from the WCF. The construct *learning* appears to be somewhat elusive in terms of its operationalization. For example, references are made to writing development (Truscott & Hsu, 2008), writing performance (Aghajanloo, Mobini, & Khosravi, 2016), written or writing accuracy (van Beuningen et al., 2012), increased accuracy (Shintani et al., 2014), formal accuracy (Ashwell, 2000), reduction of errors and so on. Indeed, the crucial inconsistency in several studies' designs and resulting analyses of the data lies in the fact that grammatical accuracy of specific linguistic items is usually not controlled and learning is typically subsumed in an overall report of several aspects of grammatical knowledge or global accuracy rate (Bruton, 2009).

Third, and perhaps crucial from an ISLA research perspective, is the failure of WCF studies, with two notable exceptions (Caras, 2019; Amelohina, Nicolás-Conesa, & Manchón, Chapter 7, present volume), to situate the research design within a language curriculum. For example, Caras (2019) followed the existing syllabus (after some initial modification was made prior to the start of the semester to accommodate the two target items) and used assigned compositions (whose topics were carefully designed to elicit the use of the target items) as her experimental tasks. While she followed the usual unfocused feedback procedure of the language curriculum to provide WCF to her participants, her study was on focused feedback on two linguistic items (the dichotomies between the Spanish copulas *ser* vs. *estar* and imperfect vs. preterit) contained within these compositions. Learning was measured in relation to participants' performances on the two linguistic items under investigation employing a pretest-immediate posttest-delayed posttest design and, as reported above, their attention to and subsequent (depth of) processing of the WCF was established via concurrent think aloud protocols. She reported that participants who received unfocused direct or metalinguistic WCF processed at high, medium, and low levels while participants who received indirect WCF processed primarily at a low level. For the copula, participants who received direct WCF outperformed significantly those in the other experimental conditions (metalinguistic, indirect, and control) while the metalinguistic WCF group also outperformed the control group at the time of Draft 2. However, these superior performances were not maintained two weeks later. For the preterit versus imperfect, type of WCF did not appear to play a significant role in subsequent performances on both the immediate and delayed posttests.

Amelohina et al. (Chapter 7, this volume) also situated their task repetition study with the aid of WCF study within a specialized 6-month program designed to prepare students for taking the ISE II exam (B2 level according to the Common European reference Framework) at the end of this course. During the duration of the course, participants were provided with either direct or indirect feedback (counter-balanced) on their four compositions, each based on a previous article or report, and then requested to rewrite each composition. Learning was measured by the complexity, accuracy, and fluency (CAF) demonstrated on the compositions. To operationalize noticing of the errors, participants were requested to write in a noticing table their errors (grammar, lexis, spelling, or punctuation) and the reason for their errors. They reported linear and non-linear effects of task repetition on diverse components of CAF, as well as a differential appropriation of indirect WCF across time, which is extremely relevant from a pedagogical perspective.

In summary, the strand of *writing-to-learn* in relation to potential L2 development via WCF is beginning to reconsider the use of one-shot designs and move toward deeper probing into L2 writers' processing and the cognitive processes employed during the writing process (see Manchón & Leow, Chapter 14, present volume, for further elaboration). In addition, the important need to situate future studies within the instructed setting as evidenced by the two recent studies (Caras, 2019; Amelohina et al., Chapter 7, this volume) clearly augurs well for a better understanding of this strand of research, especially in light of providing language teachers with robust pedagogical ramifications. The next section addresses the need to investigate this *writing-to-learn* strand of research from a curricular approach.

A curricular approach and future directions

As observed above, although L2 writing research, and more specifically, the ISLA-oriented perspective of *writing-to-learn*, has undergone tremendous growth over the past two decades, very little of the research has been conducted in relation to the environment within which it is situated, namely, within the language curriculum. Polio (2012) appeared to hint at this curricular perspective when she wrote "[W]ritten error correction is probably the most time-consuming practice teachers use, and thus worth investigating at a practical level, even without reference to specific theories." (p. 376). Manchón (2012) and Manchón and Vasylets (2019) propose the *writing-to-learn* strand of research to include both L2 learners and instructors. Research addressing both L2 writer (Li, 2017) and teacher (Gurzynski-Weiss, 2017) characteristics is clearly warranted. Similarly, Leow (2019a, b) and Leow and Cerezo (2016)) directly call for ISLA research to conduct research within the confines of

the curriculum, closely aligned to the syllabus and learning outcomes, and studied over longer periods of time to simulate the instructed learning environment.

A curricular approach easily reveals that the writing component of any language curriculum holds much value in relation to the type and amount of feedback provided to L2 learners in an instructed setting. In the high school environment (not unlike many L2 classes at higher academic levels), Harklau (2002) observed the disparity of production during face-to-face interaction in relation to L2 development and written output that was "far more copious and varied, ranging from word or phrase worksheet response to multipage multiple drafts of essays" (p. 331). Similarly, a curricular approach also underscores that the writing component in a language curriculum does not exist in isolation but is typically integrated with other skills. It is not unusual for certain target items in WCF to be covered within other activities or tasks in the classroom or even in the textbooks. To this end, future research may want to control this potential "contamination" and/or address this potential integration with other skills in relation to type of error, language proficiency, amount of exposure and so on.

This approach also demands that researchers need to address the dearth of concurrent data on the cognitive processes employed by L2 writers at all levels of proficiency during both the composing and revision phases of the writing process. Understanding how L2 writers initially compose and then interact with WCF during these two phases and what role their cognitive processes play in subsequent L2 development in relation to WCF will allow researchers to avoid making assumptions on how L2 writers process the L2 data or WCF and clearly lead to a better understanding of the *writing-to-learn* process. Concurrent data will also provide deep insights into, for example, how L2 writers process different types of linguistic items, perhaps based on the characteristics of the error produced (e.g., saliency, complexity etc.), the role of depth of processing (e.g., Caras, 2019; Cerezo et al., 2019; Leow, 2015; Manchón & Williams, 2016; Manchón & Roca de Larios, 2007; Manchón et al., 2020) during both the writing and revision process and so on. With this concurrent focus, previous WCF research that has addressed a multitude of variables should be replicated with their research designs situated within the language classroom and curriculum. Ultimately, such rich data can only lead to pedagogical implications aimed at promoting robust learning from the ISLA *writing-to-learn* strand of research.

Finally, based on students' learning outcomes, and in line with what is expected in a language classroom, researchers need to avoid using in future studies global scores (e.g., number of errors per 100 words, number of error-free T-units, etc.) and analyze L2 writers' actual errors produced on compositions (Bruton, 2009) submitted during the semester. Together with concurrent data on their processes,

researchers should be able to isolate patterns across individual L2 writers. In turn, these insights may be used to maximize the writing and revision process of students with the potential of promoting better learning. This approach supports previous calls for more qualitative analyses and, more importantly and from a curricular perspective, performed over a period of time (e.g., van Beuningen, 2010; Bruton, 2009; Storch, 2010; Storch & Wigglesworth, 2010). (See Manchón & Leow, Chapter 14, present volume, for further methodological elaboration).

Conclusion

Based on the educational level (e.g., high school, college etc.), methodology, curricular learning outcomes and largely influenced by teachers' perceptions of language learning and teaching, the amount of time spent producing the L2 may not only vary tremendously but potentially skew toward written over oral production across all language levels. In turn, the provision of feedback, and especially on a personal level, is more substantial in the written mode when compared to the oral mode. To this end, there is a clear motivation to probe deeper into the potential impact the writing component in the L2 curriculum may have on L2 development.

To add to and advance the ISLA research agenda for the *writing-to-learn* strand, this chapter has argued for a three-prong approach that is theoretically-driven (process-oriented), empirically supported, and curricular. Situating the *writing-to-learn* strand of research within a process-oriented and curricular approach is of paramount importance if researchers would like to extrapolate their empirical findings to the instructed setting. Such a three-prong approach clearly has major pedagogical implications given that (1) the data gathered from the L2 writer, both online and offline, are produced authentically (versus within a laboratory-based setting) within a given syllabus and over a period of time to simulate the language curriculum; (2) both the teacher and the L2 writer are involved in the process; (3) there is ecological validity in the findings; and (4) teachers are more prone to regard the findings as pertinent to curricular learning outcomes. Arriving at a deeper understanding, via concurrent data, of *how* L2 writers compose and process WCF within a language curriculum is clearly challenging but with huge benefits in promoting robust learning in the instructed setting from a *writing-to-learn* perspective. This research direction easily aligns well with Cumming's (Chapter 2, this volume) direction regarding future research agendas that are centrally focused on "L2 learning *through* writing".

References

Adrada-Rafael, S., & Filgueras-Gómez, M. (2019). Reactivity, language of think aloud protocol, and depth of processing in the processing of reformulated feedback. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 201–213). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-14

Aghajanloo, K., Mobini, F., & Khosravi, R. (2016). The effect of teachers' written corrective feedback (WCF) types on intermediate EFL learners' writing performance. *Advances in Language and Literary Studies*, 7, 28–37.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227–257. https://doi.org/10.1016/S1060-3743(00)00027-8

Benson, S., & DeKeyser, R. (2019). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research*, 23(6), 702–726. https://doi.org/10.1177/136216881870921

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17, 102–118. https://doi.org/10.1016/j.jslw.2007.11.004

Bitchener, J. (2016). To what extent has the published written CF research aided our understanding of its potential for L2 development? *ITL – International Journal of Applied Linguistics*, 167, 111–131. https://doi.org/10.1075/itl.167.2.01bit

Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. London: Routledge. https://doi.org/10.4324/9780203832400

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12, 409–431. https://doi.org/10.1177/1362168808089924

Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31, 193–214. https://doi.org/10.1093/applin/amp016

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781783095056

Bruton, A. (2009). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing*, 18, 136–140. https://doi.org/10.1016/j.jslw.2009.02.005

Caras, A. (2019). Written corrective feedback in compositions and the role of depth of processing. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 188–200). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-13

Cerezo, L., Manchón, R. M., & Nicolás-Conesa, F. (2019). What do learners notice while processing written corrective feedback? A look at depth of processing via written languaging. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 173–187). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-12

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of second language writing*, 12(3), 267–296. https://doi.org/10.1016/S1060-3743(03)00038-9

Coyle, Y., Cánovas-Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, 42, 25–43. https://doi.org/10.1016/j.jslw.2018.09.002

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511. https://doi.org/10.1177/0741088390007004003

DeKeyser, R. M. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd ed., pp. 94–112). London: Routledge.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT journal*, 63, 97–107. https://doi.org/10.1093/elt/ccn023

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36, 353–371. https://doi.org/10.1016/j.system.2008.02.001

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178–190). Cambridge: CUP. https://doi.org/10.1017/CBO9781139524551.016

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1–10. https://doi.org/10.1016/S1060-3743(99)80110-6

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge: CUP. https://doi.org/10.1017/CBO9781139524742.007

Ferris, D. R., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10, 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22, 307–329. https://doi.org/10.1016/j.jslw.2012.09.009

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in L2 vocabulary acquisition by means of eye tracking. *Studies in Second Language Acquisition*, 35, 483–517. https://doi.org/10.1017/S0272263113000119

Gurzynski-Weiss, L. (2017). L2 instructor individual characteristics. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 451–467). New York, NY: Routledge. https://doi.org/10.4324/9781315676968-25

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11, 329–350. https://doi.org/10.1016/S1060-3743(02)00091-7

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39, 83–101. https://doi.org/10.1017/S0261444806003399

Karimi, S. H. (2016). Effects of different types of teacher written corrective feedback on Iranian EFL learners' writing accuracy. *Journal of Applied Linguistics and Language Research*, 3(2), 216–229.

Kassim, A., & Ng, L. L. (2014). Investigating the efficacy of focused and unfocused corrective feedback on the accurate use of prepositions in written work. *English Language Teaching*, 7, 119–130. https://doi.org/10.5539/elt.v7n2p119

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal*, 7, 305–313. https://doi.org/10.1111/j.1540-4781.1991.tb05359.x

Lalande, J. F. (1982). Reducing composition errors: An experiment. *Modern Language Journal*, 66, 140–149. https://doi.org/10.1111/j.1540-4781.1982.tb06973.x

Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In R. DeKeyser (Ed.), *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 111–137). Cambridge: CUP. https://doi.org/10.1017/CBO9780511667275.007

Leow, R. P. (1999). The role of attention in second/foreign language classroom research: Methodological issues. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic Linguistics: Papers from the 2nd. Hispanic Linguistics Symposium* (pp. 60–71). Somerville, MA: Cascadilla Press.

Leow, R. P. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania*, 84, 496–509. https://doi.org/10.2307/3657810

Leow, R. P. (2012). Explicit and implicit learning in the L2 classroom: What does the research suggest? *The European Journal of Applied Linguistics and TEFL*, 2, 117–129.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge. https://doi.org/10.4324/9781315887074

Leow, R. P. (2019a). ISLA: How implicit or how explicit should it be? Theoretical, empirical, and pedagogical/curricular issues. *Language Teaching Research*, 23, 476–493.
https://doi.org/10.1177/1362168818776674

Leow, R. P. (2019b). From SLA > ISLA > ILL: A curricular perspective. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 485–493). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. P., & Cerezo, L. (2016). Deconstructing the "I" and "SLA" in ISLA: One curricular approach. *Studies in Second Language Learning and Teaching*, 6, 43–63.
https://doi.org/10.14746/ssllt.2016.6.1.3

Leow, R. P., & Martín, A. (2017). Enhancing the input to promote salience of the L2: A critical overview. In S. M. Gass, P. Spinner, & J. Behney (Eds.) *Salience in SLA* (pp. 167–186). New York, NY: Routledge.

Li, S. (2017). Cognitive differences and ISLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 396–417). New York, NY: Routledge.
https://doi.org/10.4324/9781315676968-22

Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 15–41). Cambridge: CUP.

López, M. B., Van Steendam, E., Speelman, D., & Buyse, K. (2018). The differential effects of comprehensive feedback forms in the second language writing class: Comprehensive feedback in the L2 writing class. *Language Learning*, 68, 813–850. https://doi.org/10.1111/lang.12295

Manchón, R. M. (2009). *Writing in foreign language contexts: Learning, teaching, and research*. Clevedon, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859

Manchón, R. M. (2011a). The language learning potential of writing in foreign language contexts: Lessons from research. In T. Cimasko & M. Reichelt (Eds.), *Foreign language writing instruction: Principles and practices* (pp. 44–64). Anderson, SC: Parlor Press.

Manchón, R. M. (Ed.). (2011b). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31

Manchón, R. M. (2011c). Situating the learning-to-write and writing-to-learn dimensions of L2 writing. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 3–14). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.03man

Manchón, R. M. (2012). *L2 writing development: Multiple perspectives*. Berlin: De Gruyter Mouton.
https://doi.org/10.1515/9781934078303

Manchón, R. M., Nicolás-Conesa, F., Cerezo, L. & Criado, R. (2020). L2 writers' processing of written corrective feedback: Depth of processing via written languaging. In W. Suzuki & N. Storch (Eds.), *Languaging in language learning and teaching* (pp. 241–265). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.55.11man

Manchón, R. M., & Roca de Larios, J. (2007). Writing-to-learn in instructed language learning contexts. In E. A. Soler & M. P. S. Jordá (Eds.), *Intercultural language use and language learning* (pp. 101–121). Berlin: Springer. https://doi.org/10.1007/978-1-4020-5639-0_6

Manchón, R. M., Roca de Larios, J. & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing: Implications for theory. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 102–129). Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859-008

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. B. Schwieter & A. Benati (Eds.), *The Cambridge Handbook of language learning* (pp. 341–362). Cambridge: CUP. https://doi.org/10.1017/9781108333603.015

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Ortega, L. (2011). Reflections on the learning-to-write and writing-to-learn dimensions of second language writing. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 237–250). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.16ort

Park, E. S., & Kim, O. Y. (2019). Learners' use of indirect written corrective feedback: Depth of processing and self-correction. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 214–228). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-15

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21, 375–389. https://doi.org/10.1016/j.jslw.2012.09.004

Polio, C., & Lee, J. (2017). Written language learning. In S. Loewen & M. Sato (Eds.), *Routledge handbook of instructed second language acquisition* (pp. 299–317). New York, NY: Routledge. https://doi.org/10.4324/9781315676968-17

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20, 83–95. https://doi.org/10.2307/3586390

Roca de Larios, J., Coyle, Y., & Nicolás-Conesa, F. (2016). Focus on writers: Processes and strategies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 267–286). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-015

Roca de Larios, J., Murphy, L., Manchón, R. M., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, 17, 30–47. https://doi.org/10.1016/j.jslw.2007.08.005

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158. https://doi.org/10.1093/applin/11.2.129

Semke, H. D. (1984). Effects of the red pen. *Foreign Language Annals*, 17, 195–202. https://doi.org/10.1111/j.1944-9720.1984.tb01727.x

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41, 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Shepherd, R. P., Daily O'Meara, K., Snyder, S. E. (2016). Grammar agreements: Crafting a more finely tuned approach to corrective feedback. *Journal of Response to Writing*, 2(1), 43–57.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, 22, 286–306. https://doi.org/10.1016/j.jslw.2013.03.011

Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' understanding and use of two English grammatical structures. *Language Learning*, 64(1),103–131. https://doi.org/10.1111/lang.12029

Stefanou, C., & Revesz, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *Modern Language Journal*, 99, 263–282. https://doi.org/10.1111/modl.12212

Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies*, 10, 29–46. https://doi.org/10.6018/ijes/2010/2/119181

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake and retention of corrective feedback on writing. *Studies in Second Language Acquisition*, 32, 303–334. https://doi.org/10.1017/S0272263109990532

Suzuki, W. (2012). Written languaging, direct correction, and second language writing revision. *Language Learning*, 62, 1110–1133. https://doi.org/10.1111/j.1467-9922.2012.00720.x

Suzuki, W. (2017). The effect of quality of written languaging on second language learning. *Writing & Pedagogy*, 8, 461–482. https://doi.org/10.1558/wap.27291

Suzuki, W. & N. Storch (Eds.). (2020). *Languaging in language learning and teaching*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.55

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (2005). The Output Hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371–391. https://doi.org/10.1093/applin/16.3.371

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, 8(2), 111–122. https://doi.org/10.1016/S1060-3743(99)80124-6

Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, 13, 337–343. https://doi.org/10.1016/j.jslw.2004.05.002

Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, 17, 292–305. https://doi.org/10.1016/j.jslw.2008.05.003

Van Beuningen, C. G. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies*, 10, 1–27. https://doi.org/10.6018/ijes/2010/2/119171

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learners' written accuracy. *ITL-International Journal of Applied Linguistics*, 156, 279–296. https://doi.org/10.2143/ITL.156.0.2034439

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62, 1–41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Wang, T., & Jiang, L. (2015). Studies on written corrective feedback: Theoretical perspectives, empirical evidence, and future directions. *English Language Teaching*, 8, 110–120.

Xu, C. (2009). Overgeneralization from a narrow focus: A response to Ellis et al. (2008) and Bitchener (2008). *Journal of Second Language Writing*, 18, 270–275. https://doi.org/10.1016/j.jslw.2009.05.005

Advances in empirical research

The effects of task repetition across modalities and proficiency levels

Alberto J. Sánchez[1], Rosa M. Manchón[1] and Roger Gilabert[2]
[1]University of Murcia / [2]University of Barcelona

Research on task repetition (TR) has consistently showed beneficial effects for L2 oral production in terms of CAF measures (complexity, accuracy, and fluency). However, open questions exist regarding the modality-dependency and proficiency-dependency of TR learning affordances. To advance research in these two areas, the study reported in this chapter investigated TR across modalities and L2 proficiency levels. The participants were 29 (16 high-school and 13 university) Spanish EFL learners who completed a decision-making task twice, either orally or in writing. Their task performance was assessed in terms of CAF measures using a wide range of indices. Results confirm some of the previous predictions regarding the modality-dependency of TR effects, and additional modality-related effects were also found. These findings shed light on the greater language learning affordances of writing as compared to those created by speaking tasks.

Introduction

A 'task' is considered as a "rich and potentially powerful construct" whose most relevant role is "to provide a meaningful context for language use and hence for learning" (Bygate, 2018a, p. 3). Since its beginnings, Task-Based Language Teaching (TBLT) research has explored diverse task implementation variables with the aim of creating the meaningful context for language use and language learning mentioned by Bygate. One of the variables that has attracted growing scholarly attention is task repetition (TR), recently defined as "repetition of a given configuration of purposes, and a set of content information" (Bygate, 2018a, p. 1). This configuration of variables had been previously acknowledged by Ellis (2012), who suggested that TR could be implemented "without any changes to the task, or by modifying the design of the task or by manipulating one of the other implementation variables" (p. 202).

Task repetition and language learning: Rationale

The theoretical and applied relevance of research in the domain derives from the consideration of the learning that may result from repeating speaking tasks. The rationale for such potential learning is both psycholinguistic and pedagogical. From the first angle, the connection between TR and language learning is linked to the attentional demands associated with the implementation of the components of speech production (i.e. conceptualization, formulation, and articulation, Levelt, 1989). Thus, as speaking takes place in real time, L2 users have to divide their attentional resources among creating a message (conceptualisation), retrieving the language needed to create adequate meaning-form mappings to express their intended meaning (formulation), and, finally, uttering the message in order to convey his/her own ideas (articulation) under the time constrains that characterise oral communication in real time. These different processes compete for attention and, therefore, strategic decisions need to be taken regarding which process to prioritize and when. On the basis of these assumptions, Bygate (2001) claimed that when learners repeat a task, they "draw on the conceptual structuring of the information and on encodings which they have previously used" (p. 253), thus reducing the cognitive demands of the task and freeing up attentional resources that may be devoted to the formal aspect of the task. This new "opportunity […] to rework their language" (Bygate & Samuda, 2005, p. 114) has been claimed to promote language development (Bygate, 1996, 2001, 2006, 2018a, b).

Therefore, the learning affordances of TR are premised on the constraints on attentional resources that result from the real-time, on-line nature of oral communication. Yet, given the off-line nature of most forms of written communication, a relevant empirical question in the domain is whether or not the purported language learning gains that may derive from TR are modality-dependent. In other words, given that writers have more ample time-on-task, it could be hypothesized (see Manchón, 2014a, b) that, in principle, they are more likely to be able to divide their attentional resources among various task demands concurrently, hence succeeding in focusing on meaning and language simultaneously. It has also been claimed that the recursive, problem-solving nature of writing, along with the permanence and visibility of the written text, also facilitate the learners' engagement in focus on form processes during the entire task (Manchón, 2014a, b; Manchón & Williams, 2016; Williams, 2012), something that is more difficult or even impossible to achieve under the constraints of real-time oral communication, precisely the type of communication that has featured more prominently in TR research (see reviews in Manchón, 2014a; Nitta & Baba, 2018). In short, the central argument of the TR literature in speaking, i.e., the benefits of repeating a task because of attentional demands, may apply differently to the performance of writing tasks.

One further consideration in connection with a possible modality-dependency of TR learning effects is worth mentioning: Distinctive TR effects across modalities need to take into account the role of planning in task performance. In this sense, the first iteration of an oral task is thought to function as a planning condition (Ahmadian, 2011; Ellis, 2015) leading to increased (more complex, more accurate and/or more fluent) performance in the second iteration. In contrast, writers may be able to plan during the first task iteration since greater time is available. In fact, recent accounts of L2 writing as a site for language learning (cf. Manchón, 2011a, 2014b; Williams, 2012) systematically emphasize that the off-line nature of writing allows L2 writers to devote more time and attentional resources to task conceptualization, task planning, and task completion, three processes closely associated with attention to language-related concerns.

Taken together, the above arguments point to the relevance of looking more deeply into the learning affordances of TR across modalities, as well as of putting predictions on the distinctive affordances of TR in writing to the empirical test, two aims that we pursued in the study reported in this chapter. In the next section we review the relevant empirical literature to better situate our own study.

Empirical research on task repetition in speaking and writing

A substantial body of empirical work has been devoted to ascertaining the learning that may derive from TR. Most of these studies have focused on oral tasks and the results obtained distinctively and consistently show that TR in the oral mode leads to better performance in terms of CAF (complexity, accuracy and fluency) measures although the results vary across studies (see review in Manchón, 2014a, and more recent research in contributions to Bygate, 2018b). For instance, several studies (Birjandi & Ahangori, 2008; Bygate, 1996, 2001) have reported improvements in fluency and complexity, whereas accuracy improvements were either non-existent or not statistically significant. This contrasts with other studies that have reported no improvements in fluency (Matsumara, Kawamura, & Affricano, 2008) although they did find an effect on accuracy (found to be task-dependent). Yet, the overall picture is that TR results in increases in fluency (Ahmadian, 2011; Ahmadian & Tavakoli, 2010; Bygate, 2001; Gass, Mackey, Alvarez-Torres, & Fernández-García, 1999; Hu, 2018; Kobayashi & Kobayashi, 2018; Lynch & Maclean, 2000, 2001; Sheppard & Ellis, 2018), accuracy (Bygate, 1996; Fukuta, 2015; Hu, 2018; Kim & Tracy-Ventura, 2013; Kobayashi & Kobayashi, 2018; Lynch & Maclean, 2000, 2001), and complexity (Ahmadian, 2011; Ahmadian & Tavakoli, 2010; Bygate, 2001; Fukuta, 2015; Gass et al., 1999; Kobayashi & Kobayashi, 2018; Sheppard & Ellis, 2018).

In contrast to this abundant research, TR in writing has received much less attention. Symptomatic of this trend is the fact that the most recent collection of

TR studies (Bygate, 2018b) includes just one single writing study. The few writing TR studies available (Amiryousefi, 2016; Amelohina et al., Chapter 7 this volume; Nitta & Baba, 2014, 2018) once again provide empirical confirmation of language learning benefits resulting from repeating a writing task. Thus, both Nitta and Baba (2014) and Amiryousefi (2016) reported increased performance in terms of fluency. Regarding other areas of performance, Nitta and Baba (2014) found that TR leads to greater complexity in the long term at the expense of the initial increases in fluency, although the patterns of development show distinctive variation across individuals (Nitta & Baba, 2018). Amiryousefi (2016) reported higher accuracy resulting from TR although complexity was not affected.

The divergent findings on potential effects of TR have led Bygate (2018b) to claim that "although learners' language is likely to change across iterations, we cannot confidently anticipate whether this will occur predominately in terms of complexity, or accuracy, or fluency" (p. 8). We could add that, owing to the scant research conducted to date on TR in writing, together with the disparity of empirical findings available in the oral domain, further empirical validation of predictions, as well as further expansion of the available body of empirical work, are fully justified. Also, such disparity of results may be well due to the fact that CAF measures have been expressed in different indices across studies both in speaking and in writing. Given that results could well depend on the specific CAF measures used, our research employed the most prevalent CAF measures used in the previous literature.

Adding to the gaps in research mentioned so far, the relative effectiveness and purported distinctiveness of task repetition across modalities is also under researched. This is an important gap in language learning studies from theoretical, empirical, and applied perspectives. Theoretically and empirically, given the above arguments on the modality-dependency of TR effects as a function of the timed nature of oral and written forms of communication, TR modality studies would therefore contribute to advancing TBLT theoretical work, on the one hand, and theoretical tenets on the language learning potential of writing (cf. Manchón, 2011a, b; Manchón & Williams, 2016; Williams, 2012), on the other. From an applied angle, the findings of TR modality studies regarding which modality leads to what learning gains could inform TBLT-oriented pedagogical approaches, especially regarding choice of tasks.

These considerations motivated our interest in looking into the effects of TR in speaking and writing in one single study. We further investigated whether any potential effects were mediated by the learners' L2 proficiency, a research aim once again motivated by a consideration of previous empirical work, as detailed in the next section.

The role of proficiency

In addition to modality, a number of additional task-related and learner-related factors have been claimed to mediate TR effects. Among the former, timing of the repetition (de Jong & Tillman, 2018), task complexity (Kim, Crossley, Jung, Kyle, & Kang, 2018), and task type (see Birjandi & Ahangori, 2008; Matsumara, Kawamura, & Africano, 2008; Nitta & Baba, 2014, 2018. But see Bygate, 2001, and Gass et al., 1999 for a lack of effect of task type) have been predicted or shown to influence TR outcomes. Among the learner-related factors investigated (see several contributions to Bygate, 2018b), L2 proficiency is the one that has attracted the most attention and one that is particularly relevant given the divergent theoretical predictions made and the contradictory results of extant research. Thus, whereas Ellis's (2005) claim is that TR effects would be proficiency-dependent, Bygate's (2018a) prediction is that TR is likely to function in the same ways across proficiency levels. In support of a proficiency-dependency of TR effects, it has been argued that when learners engage in TR, they count on a new opportunity to monitor their previous production. For that purpose, they may reflect on their explicit knowledge, at least partially (Ellis, 2005). As more proficient learners have a more developed L2 system than lower proficiency learners, they are more likely to take greater advantage of TR. In support of this prediction, Mojavezi (2013) found that higher proficiency participants were indeed able to take further advantage of TR in an oral narrative task in terms of fluency, syntactic complexity, and accuracy. In contrast, citing various studies that included participants at different proficiency levels (Bygate, 2001; Lynch & Maclean, 2000, 2001), Bygate (2018a) has more recently argued that "there are grounds for tentatively concluding that changes across iterations in various aspects of language are likely to arise irrespective of proficiency level" (p. 8).

These divergent predictions are still empirical questions. Further arguments for conducting research on the connection between TR learning effects and proficiency derive from, first, the consideration of the scant TR research investigating diverse proficiency levels within one single study, and, second, the fact that L2 proficiency has not been an independent variable in the extant research of TR in writing.

The present study: Research questions

Taking into account the above considerations, the aim of the study reported in this chapter is twofold. First, it aims to fill some of the aforementioned gaps in research by investigating task repetition across modalities and across proficiency levels. Second, in connection with the overall aims of the present volume, the study intends to contribute to novel insights into the language learning potential of L2

writing by contrasting learning effects across modalities, hence complementing other task-modality chapters in the present volume (see chapters 8 and 9).

On the basis of the research reviewed in the previous sections, we formulated the following research questions:

RQ1.   Does task repetition across modalities (oral, writing) result in any quantitative differences in learner's performance in terms of complexity, accuracy and fluency (CAF) measures?

RQ2.   Are any potential observed differences mediated by proficiency?

Method

Participants

The study was conducted with 29 Spanish EFL participants (12 male, 17 female), including high-school ($n = 16$) and university ($n = 13$) students representing different proficiency levels according to the Oxford Placement Test: Low (A2 and A2+ level) for high-school participants and High (B2 level and above) for university participants. No specific test comparing the groups was used. It was assumed that difference in proficiency was distinct enough provided that there was one level (B1) existing between the two groups. Table 1 shows the allocation of participants in the two proficiency groups to the two treatment conditions (TR in speaking or writing). Participants' age ranged from 14–16 and 18–19 for low-proficiency participants and high-proficiency, respectively.

Table 1. Participants' grouping

Task condition	Proficiency level	
	High (University)	Low (High school)
TR_Speaking	$n = 6$	$n = 8$
TR_Writing	$n = 7$	$n = 8$

Tasks and procedures

The Firechief task used (see Appendix) was taken from Gilabert (2005, 2007). In this decision-making task, participants were presented a picture prompt showing a building in flames in which different people needed to be rescued. To complete the task, participants had to indicate the actions they would take in order to save each of the people in the picture with the resources shown in it, to determine the

sequence of their actions, and to justify their decisions. Following the original task instructions (Gilabert, 2005, 2007), participants were given 30 seconds to familiarise themselves with the picture and task instructions. At time 1 (first iteration of the task), they were handed a coloured copy of the picture prompt and the tasks instructions, which were written in their mother tongue (Spanish) to ensure full understanding. Once they familiarized themselves with the task and task instructions, they were asked to start performing the task in speaking or in writing. Eight days later (Time 2), they were asked to repeat the same task in the same modality and under the same circumstances as during the first iteration. The speaking and writing groups performed the task individually in an empty room and their performance was digitally recorded.

Data analyses procedures

Data (transcribed oral production and written texts) were analysed in terms of those CAF measures more widely used in previous research, which were applied across modalities and proficiency levels.

Complexity was operationalised as both lexical and syntactic complexity. To gauge development in the area of lexical complexity, and on account of its multidimensional nature, the sub-constructs of lexical variety, lexical richness, and lexical sophistication were analysed. Lexical variety was measured in terms of D-value (Malvern & Richards, 2002). Lexical richness was calculated by analysing Guiraud Index (Guiraud, 1954) following Bulté and Housen (2012, 2014), which accounts for both variety and number of words. Finally, lexical sophistication was calculated through Advanced Guiraud Index (Daller, van Hout, & Treffer-Daller, 2003). RANGE software (Heatley, Nation, & Coxhead, 2002) was used to analyse this dimension. This software calculates the proportion of words that fit into different frequency word bands. Words were considered sophisticated if they were not part of the first frequency band (most frequent 1.000 words) to capture any changes in the lexical sophistication of our low-proficiency participants.

Following suggestions to analyse syntactic complexity from a general, clausal and phrasal perspectives (Norris & Ortega, 2009), five different measures were used: an overall measure of complexity i.e. the mean length of t-unit (MLT), a subordination measure i.e. dependent clauses/t-unit (DC/T), and a measure of coordination i.e. t-units/sentence (T/S). The measure of coordination was included due to previous claims about its relevance when dealing with beginner learners (Bardovi-Harling, 1992). These three dimensions were analysed using the Web-based L2 Syntactical Complexity Analyzer (Lu, 2010, 2011; Lu & Ai, 2015). Additionally, two other syntactic complexity measures were targeted by using the online software Coh-Metrix 3.0 (McNamara, Graesser, McCarthy, & Cai, 2014), namely noun phrase complexity

(modifiers/NP) and an innovative measure (STRUTt – sentence syntax similarity), hardly used in previous research (see Nitta & Baba, 2014; Mazgutova & Kormos, 2015 for notable exceptions), operationalised as syntactic variety, and regarded as an integral part of syntactic complexity (Bulté & Housen, 2012). STRUTt is a reversely-coded measure as it reflects sentence syntax similarity. The higher the score, the more similar, i.e. the least varied, syntax is. Therefore, a high score in this measure implies a negative effect on this area of syntactic complexity. Although few studies have brought together both syntactic complexity and variety, syntactic complexity needs to be approached multidimensionally, hence the reason to include such a measure.

Accuracy was analysed through measures that consider total number of errors, a decision based on the unlikelihood of finding error-free t-units when dealing with low-proficiency learners (Kuiken & Vedder, 2007). Bardovi-Harling and Bofman's categories (1989) were adapted, differentiating between morpho-syntactic errors and lexical errors. The measures selected were total errors per 100 words and total errors per T-unit, morpho-syntactic, lexical errors per 100 words, and morpho-syntactic and lexical errors per T-unit. Errors were manually coded using CLAN (**C**omputerised **L**anguage **AN**alisis) program, part of Child Language Database Exchange System (CHILDES). Regarding reliability scores, a simple intra-rater and inter-rater approach based on percentage agreement was used. Within a one-week time interval, the coder re-analysed 13,79% of the data applying the same error categories i.e. lexical errors and morpho-syntactic errors. Intra-rater reliability scores reached 100%. Additionally, a second independent researcher coded 13,79% of the data applying again the same error categories. Inter-coder agreement scores reached 100% in every error category coded, i.e. lexical errors and morpho-syntactic errors. Whenever disagreement arose, the two researchers discussed each categorization until complete agreement was reached.

Speed fluency (Skehan, 2009) was operationalized in terms of both words per minute (total number of words over total number of task-time, measured in minutes) and syllables per minute (total number of syllables over the total number of task-time, measured in minutes).

Data analysis procedures

Following Plonsky (2015), we used descriptive statistics (means and standard deviations -SDs), effects sizes (η^2) and confidence intervals (CIs) for several reasons, especially our small sample size. Descriptive statistics, effect sizes and CIs were calculated using SPSS. Cohen's guidelines (1988) were followed to measure effect sizes and were considered small ($\eta^2 = .01$), medium ($\eta^2 = .06$) or large ($\eta^2 = .14$). All the effects reported in the next section are considered large effects.

Results

The two questions guiding our study asked about the modality-dependency and proficiency-dependency of TR in terms of CAF measures. In what follows, results will be presented according to the different dependent variables, i.e. CAF measures, and with respect to the effects found across modalities and proficiency levels.

Complexity

Syntactic complexity. Table 2 shows the descriptive statistics for syntactic complexity measures for the oral and writing groups at times 1 and 2. As can be seen in Table 3, which shows the mean score at times 1 and 2 for all groups, 95% confidence intervals and effect sizes, a large effect was found in the measure for subordination i.e. dependent clauses per t-unit (DC/T), ($\eta^2 = .18$). Table 3 shows that the mean score

Table 2. Means and standard deviations for syntactic complexity measures across groups

Measures		Low proficiency		High proficiency	
		Oral group (*n* = 8)	Writing group (*n* = 8)	Oral group (*n* = 6)	Writing group (*n* = 7)
		M (SD)	M (SD)	M (SD)	M (SD)
T1	MLT	150.29 (71.64)	180.85 (52.99)	173.05 (25.77)	187.36 (50.04)
	DC/T	2.85 (4.42)	4.33 (5.52)	4.40 (5.89)	5.43 (6.16)
	T/S	13.10 (6.38)	7.77 (6.17)	12.30 (1.61)	11.09 (5.04)
	NPC	.73 (.19)	.91 (.19)	.78 (.13)	.86 (.07)
	STRUTt	.11 (.08)	.16 (.14)	.11 (.04)	.10 (.03)
T2	MLT	148.25 (53.01)	157.3 (61.40)	148.25 (53.02)	173.67 (19.20)
	DC/T	2.47 (5.47)	1.79 (3.8/3)	2.67 (4.93)	2.03 (3.52)
	T/S	13.37 (2.21)	8.91 (5.36)	12.05 (1.60)	12.80 (1.79)
	NPC	.78 (.13)	.95 (.23)	.72 (.13)	.89 (.15)
	STRUTt	.13 (.04)	.15 (.06)	.13 (.05)	.09 (.02)

MLT: mean length of t-unit. **DC/T**: dependent clauses per t-unit. **T/S**: t-units per sentence. **NPC**: noun phrase complexity. **STRUTt**: sentence syntax similarity – syntactic variety.

Table 3. 95% confidence intervals and effect size for syntactic complexity (within-subject)

Measure	Time	Mean	95% confidence interval		Effect size η^2
			Lower bound	Upper bound	
DC/T	1	4.25	2.14	6.36	.18
	2	2.24	.51	3.98	

DC/T: dependent clauses per t-unit.

for subordination (DC/T) for all groups at time 1 falls beyond the upper bound for 95% confidence intervals at time 2 for that measure. This result in effect reflects a decrease in performance for all groups from the first to the second iteration of the task, a decrease particularly noticeable in the writing group. No other large effect was observed. In sum, task repetition – across modalities and proficiency levels – did not result in any beneficial effects in the area of syntactic complexity and even had detrimental effects in the area of subordination, which was more marked in the written data.

Lexical complexity. As shown in Table 4, TR did not result in any large effect in the lexical complexity of the text written by the participants at times 1 and 2 irrespective of modality of production or proficiency level in terms of variety (D) ($\eta^2 = .01$), richness (G), ($\eta^2 = .04$), and lexical sophistication (LS) ($\eta^2 = .03$).

Table 4. Means and standard deviations for lexical complexity measures across groups

Measures		Low proficiency		High proficiency	
		Oral group (*n* = 8)	Writing group (*n* = 8)	Oral group (*n* = 6)	Writing group (*n* = 7)
		M (SD)	M (SD)	M (SD)	M (SD)
T1	D	22.35	22.54	33.75	45.24
		(6.98)	(12.43)	(9.37)	(9.65)
	Guiraud	4.60	4.01	5.56	7.18
		(.72)	(.86)	(.87)	(1.15)
	LS	.33	.34	.48	1.02
		(.13)	(.24)	(.27)	(.56)
T2	D	24.10	17.40	31.88	48.02
		(8.83)	(10.48)	(10.32)	(13.63)
	Guiraud	4.72	4.01	5.46	7.12
		(.93)	(.86)	(1.05)	(1.23)
	LS	.39	.34	.62	.99
		(.30)	(.18)	(.46)	(.51)

D: lexical variety. **G**: lexical richness. **LS**: lexical sophistication.

Modality-related effects

Other large effects were found when analysing between-subject factors, as can be seen in Tables 5 and 6, which show 95% confidence intervals and effects sizes for syntactic complexity and lexical complexity, respectively. These results indicate that the written modality prompted the use of more complex language in terms of lexical diversity, lexical richness, and lexical sophistication by high-proficiency learner, as well as more complex language in terms of noun phrase complexity across proficiency levels.

Table 5. 95% confidence intervals and effect sizes for syntactic complexity measures (between-subject factors)

Measure	Group	Mean	95% confidence interval		Effect size η^2
			Lower bound	Upper bound	
NPC	Oral	.75	.67	.83	.24
	written	.90	.83	.98	

NPC: noun phrase complexity.

Table 6. 95% confidence intervals and effect sizes for lexical complexity measures (between-subject factors)

Measure	Proficiency	Group	Mean	95% confidence interval		Effect size η^2
				Lower bound	Upper bound	
D	high	Oral	32.81	24.79	40.84	.19
		written	46.63	39.20	54.06	
LS	high	Oral	.55	.26	.83	.13
		written	1.00	.74	1.26	
G	high	Oral	5.51	4.71	6.31	.25
		written	7.15	6.41	7.89	

D: lexical variety. **G**: lexical richness. **LS**: lexical sophistication.

Thus, as seen in Table 5, a large effect was found for noun phrase complexity (NPC), (η^2 = .24). The overall mean score for writing groups falls beyond the upper bound of 95% confidence intervals for oral groups. This result indicates that both the high and low proficiency participants in the writing groups used more complex language in terms of noun phrase complexity as compared to participants in the oral groups.

As for lexical complexity, Table 6 shows the large effects found in the measures for variety (D) (η^2 = .19) and richness (G) (η^2 = .25). Regarding lexical sophistication (LS), a relevant medium effect was found (η^2 = .13), very close to being large (large effect established at η^2 = .14, (Cohen, 1988)). These large effects show that writers used more complex language than speakers in terms of both lexical diversity,

lexical richness and, although to a lesser extent, lexical sophistication. In this case, the results were mediated by proficiency as the effect was observed only in the data of the higher proficiency writers.

Accuracy

Table 7 shows the means and standard deviations for accuracy measures across groups. No noticeable effects as a function of TR were observed in any measure regardless of modality of production or proficiency level: Morpho-syntactic errors per 100 words (MSE/100w) (η^2 = .05), lexical errors per 100 words (LEXE/100w),

Table 7. Means and standard deviations for accuracy measures across groups

Measures		Low proficiency		High proficiency	
		Oral group (*n* = 8)	Writing group (*n* = 8)	Oral group (*n* = 6)	Writing group (*n* = 7)
		M (SD)	M (SD)	M (SD)	M (SD)
T1	MSE/100w	11.06 (4.42)	12.52 (5.62)	.04 (.03)	.03 (.01)
	LEXE/100w	4.40 (2.92)	4.38 (3.39)	.01 (.01)	.01 (.01)
	TOTALE/100w	15.45 (6.30)	16.90 (7.13)	.05 (.02)	.04 (.02)
	MSE/T	1.63 (.33)	1.76 (.81)	.76 (.50)	.59 (.22)
	LEXE/T	.68 (.52)	.61 (.46)	.16 (.16)	.19 (.12)
	TOTALE/T	2.31 (.77)	2.37 (1.06)	.92 (.43)	.78 (.30)
T2	MSE/100w	9.18 (3.70)	13.96 (8.81)	.03 (.02)	.02 (.01)
	LEXE/100w	6.02 (5.30)	6.60 (7.25)	.00 (.01)	.00 (.01)
	TOTALE/100w	15.21 (6.92)	20.55 (15.18)	.04 (.02)	.03 (0.1)
	MSE/T	1.38 (.50)	1.87 (1.05)	.46 (.24)	.41 (.16)
	LEXE/T	.86 (.72)	.88 (.90)	.12 (.11)	.13 (.13)
	TOTALE/T	2.23 (.94)	2.74 (1.80)	.58 (.25)	.53 (.18)

MSE/100: morpho-syntactic errors per 100 words. **LEXE/100**: lexical errors per 100 words. **TOTALE/100**: total errors per 100 words. **MSE/T**: morpho-syntactic errors per t-unit. **LEXE/T**: lexical errors per t-unit. **TOTALE/T**: total errors per t-unit.

($\eta^2 = .00$), total number of errors per 100 words (TOTALE/100w), ($\eta^2 = .03$), morpho-syntactic errors per t-unit (MSE/T), ($\eta^2 = .02$), lexical errors per t-unit (LEXE/T) ($\eta^2 = .00$), total number of errors per t-unit (TOTALE/T) ($\eta^2 = .01$).

Fluency

Table 8 shows descriptive statistics for fluency measures across groups while Table 9 shows means, 95% confidence intervals and effect sizes for oral groups. Large effects were found in the two measures of fluency (words per minute and syllables per minute), a finding that was mediated by modality: Only those participants who engaged in TR in the oral mode were more fluent during the second iteration of the task, an increase observed across proficiency levels. Thus, as shown in Table 9, large effects were found in the measures of words/minute (W/M) ($\eta^2 = .33$), and syllables/minute (S/M) ($\eta^2 = .39$).

Table 8. Means and standard deviations for fluency measures across groups

Measures		Low proficiency		High proficiency	
		Oral group (n = 8)	Writing group (n = 8)	Oral group (n = 6)	Writing group (n = 7)
		M (SD)	M (SD)	M (SD)	M (SD)
T1	W/M	39.66 (23.56)	5.18 (3.10)	90.61 (16.85)	12.76 (3.28)
	S/M	49.57 (28.48)	6.45 (3.68)	111.14 (21.82)	15.66 (5.05)
T2	W/M	54.73 (23.13)	7.55 (3.58)	112.80 (13.19)	18.59 (7.15)
	S/M	68.84 (27.50)	9.40 (4.00)	131.84 (18.73)	24.21 (9.66)

W/M: words per minute. S/M: syllables per minute.

Table 9. 95% confidence intervals and effect sizes for fluency measures

Mean	Group	Time	Mean	95% confidence interval		Effect size η^2
				Lower bound	Upper bound	
W/M	Oral	1	65.14	56.93	73.34	.33
		2	83.76	75.88	91.63	
S/M	Oral	1	80.35	70.21	90.49	.39
		2	100.34	90.57	110.11	

W/M: words per minute. S/M: syllables per minute.

Discussion

Our first research questions asked whether task repetition in the two modalities (oral/writing) would result in any quantitative differences in task outcomes as measured by CAF indices. Two main sets of findings obtained are worth discussion. First, in contrast to many other TR studies reviewed in the first section of the chapter, we found that repeating tasks in speaking or writing by higher and lower L2 proficiency learners did not result in any noticeable positive effect in the areas of lexical complexity and accuracy and even had a detrimental effect in the area of syntactic complexity, as all proficiency groups reduced their rate of global subordination from the first to the second iteration of both the speaking and the writing tasks. This reduction in syntactic complexity was more marked for participants who engaged in task repetition in writing. Second, different modality-related effects were observed. Writing elicited more complex language in terms of one dimension of syntactic complexity (noun phrase complexity) and lexical complexity (lexical richness, lexical variety and lexical sophistication), although this last effect was mediated by proficiency. Finally, the large effects found in the area of fluency may indicate that TR leads to more fluent oral performance regardless of proficiency level.

The observed benefits in the area of fluency for the oral group coincide with some findings reported in previous studies (Ahmadian, 2011; Ahmadian & Tavakoli, 2010; Bygate, 2001; Gass et al., 1999; Hu, 2018; Kobayashi & Kobayashi, 2018; Lynch & Maclean, 2000, 2001; Sheppard & Ellis, 2018) and can be taken as support of Bygate's (1996, 2001) prediction that when learners perform a task orally, they prioritise meaning over form during the first iteration, and they may shift their attention from meaning to form in a new iteration of the task. What our study adds is that these anticipated fluency effects of TR were not found to be proficiency-dependent. In contrast, TR in writing did not result in higher fluency as compared to speaking, regardless of proficiency. These findings are not in line with those in the scant research done to date on TR in the written modality (Amiryousefi, 2016; Nitta & Baba, 2014). For instance, Nitta and Baba (2014) found TR to lead to greater complexity in the long term at the expense of initial increases in fluency, a benefit of immediate TR that was not observed in our data. However, we should be cautious when comparing our results with those in previous studies given the different types of tasks used: TR has mostly been studied though narrative tasks – both orally and in writing – while we used a decision-making task. This points to the relevance of including task type as a relevant variable in future empirical research on TR effects.

As mentioned above, the detrimental effect of TR found in the area of subordination (DC/T) was more marked in the writing groups than the one observed in the spoken data. This observed reduction in syntactic complexity (which adds

to the mixed, and at times conflicting, effects of TR reported in previous research) points to the complexity that may characterize TR effects, and supports Bygate's (2018a) claim about the difficulty in anticipating distinct effects of TR. Similarly, our data seem to contradict the potential beneficial effects of TR in the writing domain anticipated by Manchón (2014a, b), who predicted that "the availability of time that characterizes writing may represent an ideal condition for TR to foster deeper linguistic processing" because this extra time could "allow L2 writers to be more in control of their attentional resources, more prone to prioritize linguistic concerns (in contrast to what is possible in oral production) and, accordingly, more likely to attend to language" (Manchón, 2014a, p. 20).

The observed reduction in syntactic complexity observed in our data would point in the opposite direction. However, the pattern is once again a complex one given the interaction between task modality and proficiency, as discussed next.

Our second research question asked about the potential proficiency-dependency of any observed TR effects. From one perspective, our data would lead to the conclusion that TR effects are not mediated by proficiency, this being one of the two opposing predictions in the field, as noted in the introductory section: The observed increase in fluency in speaking applied across proficiency levels.

Additionally, other modality-related effects were found. In the first place, we did find an effect in the writing groups in one dimension of syntactic complexity (noun phrase complexity). Both low- and high-proficiency writers used more complex language in terms of noun-phrase complexity, that is, the effects applied across proficiency levels. Secondly, an interesting interaction between modality and proficiency was also present in our data: Only high-proficiency writers were able to outperform high-proficiency speakers in terms of lexical variety, lexical richness, and lexical sophistication. Interestingly, this increase in lexical complexity was at the expense of global measures of syntactic complexity, although only for the higher proficiency group, as the texts written by the lower proficiency participants during the second iteration of the task also included less subordination.

A plausible interpretation of these findings could be made on the basis of the predicted modality-dependency of task repetition effects (cf. Manchón, 2014a, b). To recall, these claims were predicated on the temporal nature of speaking and writing and, consequently, on (i) the constraints on attentional resources to address diverse concerns simultaneously when producing oral language in real time, in contrast to (ii) what the extra time-on-task condition of writing could buy with regard to what is possible during the first iteration of the task and, as a result, the kind of benefits that could be expected during the second iteration. It could therefore be speculated that the lower proficiency participants in the writing groups in our study may have completed the task in full during their first encounter with it and, hence, they were able to make full use of their linguistic knowledge during

the first iteration. As a result, having the opportunity to engage in task repetition did not result in greater attention to linguistic concerns, hence the observed lack of positive effects of TR in any of the CAF measures used. It may be the case that for task repetition in writing to result in increased performance with low proficiency L2 users, either massed repetition (Nitta & Baba, 2014) or a combination of TR with some sort of external intervention in the form of written corrective feedback (WCF) in between tasks iterations is needed. In fact, as Manchón (2014b) has suggested, the provision of WCF may prompt learners to focus on form during the second iteration of the task, which may lead at least to increased accuracy.

In contrast, the data from the higher proficiency writing group regarding lexical complexity and the data from writing groups across proficiency levels regarding syntactic complexity serve to confirm the prediction that writing may foster the use of more complex language and adds to previous empirical work showing this effect (Ellis & Yuan, 2005; Kormos, 2014; Vasylets, Gilabert, & Manchón, 2017). As the higher proficiency writers in our study had more linguistic resources at their disposal than their lower-proficiency counterparts, and given the greater availability of time while writing, they seemed to have set more complex linguistic goals and, consequently, only these learners were able to deploy a more varied, rich, and sophisticated language. These data do partially support Manchón's (2014a, b) prediction regarding the possibility of complexification of goals during a second iteration of the task in the writing mode. What our data would suggest is that the likelihood of addressing new (and higher) language concerns is mediated by proficiency. Whether or not TR effects in writing on lexical and syntactic complexity are dependent on a given threshold level of L2 proficiency needs to be further substantiated with a wider population and a variety of tasks.

The fact that high proficiency writers in our study chose to focus their attention on improving their use of lexis across iterations of the task could also be explained from the perspective of the central role of lexical use in writing, together with the consideration of the various purposes that lexical searches can have in writing. Thus, it has been suggested that "vocabulary knowledge is central to the writing activity and, moreover, vocabulary is considered a criterion for assessing writing" (Agustín-Llach, 2011, p. 50). Furthermore, finding lexical items to express one's intended meaning has been claimed to constitute "one of the most crucial problems writers have to face" (Murphy & Roca de Larios, 2010, p. 61). Similarly, several studies (cf. Murphy & Roca de Larios, 2010; Roca de Larios, Manchón, & Murphy, 1996) have documented that lexical searches are central to the process of writing and the purpose of these searches includes both compensating for linguistic difficulties as well as attempts to improve one's lexical choices. The latter corresponds to the behavior observed for the more proficient writers in our study.

In short, globally considered, our findings point in four different relevant directions. First, they provide partial confirmation of the purported diverse beneficial effects of TR as these were found to apply only to fluency in speaking for all proficiency groups. In this sense, our data can be taken as supporting Bygate's (2018a) contention regarding the unpredicted effects of TR. We can corroborate his claim, and do so across modalities, that "although learners' language is likely to change across iterations, we cannot confidently anticipate whether this will occur predominately in terms of complexity, or accuracy, or fluency" (p. 8). Second, our findings point to distinct modality-related effects of TR: Reiterating what has just been mentioned, these modality-related TR effects were positive in the case of speaking in the area of fluency, an effect observed in the oral mode but not in writing. Third, our data showed a lack of mediation of proficiency on TR as the effects reported did not vary across proficiency levels. Given the clash of our findings with the available empirical evidence (Mojavezi, 2013), further research on the potentially proficiency-mediation of TR is needed. In the fourth place, our findings regarding higher lexical and syntactic complexity in the written mode reinforce previous assumptions on the language learning associated with written practices. Furthermore, the mediation of proficiency in modality effects appears to be much more complex than a simple dichotomous yes/no option: We found that some of the observed effects were proficiency-dependent (lexical complexity) whereas others were not (syntactic complexity), hence the relevance of further empirical research on the issue.

Finally, we would like to discuss our data form the perspective of the theme of the book: the connection between writing and language learning. We would like to suggest that the research reported in this chapter provides new empirical evidence on the language learning potential of written practices. In our study, writing elicited the use of more complex language than speaking in some dimensions of both syntactic and lexical complexity, showing at the same time a proficiency-dependency of the effects regarding lexical complexity. Again, this issue should be connected to the greater availability of time and its impact on the allocation of attentional resources while writing in a L2. This potential for deeper linguistic processing in writing than in speaking supports the view that writing may serve as a greater catalyst for the different processes involved in the learning of languages and suggests that different language modalities may foster distinct opportunities for language learning.

Conclusions

We would like to finish by reiterating the dual set of implications of our study, i.e. with respect to TR studies and to the connection between writing and language learning.

Regarding task repetition studies, our findings would support previous claims of the difficulty of anticipating precise effects of TR. In this sense, as noted in the introductory section, Bygate (2018a) has recently claimed that "although learners' language is likely to change across iterations, we cannot confidently anticipate whether this will occur predominately in terms of complexity, or accuracy, or fluency" (p. 8). Additionally, our study would support a certain role for proficiency although, much more significantly, what our data suggest is the existence of interesting patterns of interaction between task modality and proficiency. Hence, a central mandate for future research in the domain would be to disentangle this complexity. In this sense, a fruitful future research direction would be to investigate the task-dependency of TR effects, as well as the interaction between task-, modality-and proficiency factors. These interactions need to be investigated both cross-sectionally and longitudinally given the available evidence that TR effects may vary across time. In this respect, Nitta and Baba (2018), referring to their own previous research, report that "while writing task repetition did not necessarily improve their command of language immediately, over time it was likely to encourage their learning in terms of fluency […]and lexical and grammatical complexity" (p. 286). Finally, future TR research agendas must address the potential effects of WCF, and do so for both theoretical and pedagogical reasons.

From a different angle, our study can shed light on the connection between writing and language learning in terms of immediate improved use of language when repeating tasks orally and in writing. Therefore, our study has nothing to say about how writing may contribute to the consolidation or expansion of new knowledge. What our research allows us to conclude is that, first, the dimensions of language attended to when tasks are repeated vary across modalities, hence the conjecture that different language modalities may foster distinct opportunities for language learning. Second, writing elicited more complex language, especially at the lexical level, a finding that was proficiency-dependent. Therefore, another central mandate for future research on writing as a site for language learning is the study of the proficiency-dependence of the purported language learning affordances of writing. Equally relevant would be to investigate the mediation of task-related variables as there are indications that task complexity mediates the relationship between TR and vocabulary use when repeating speaking tasks (see Kim et al., 2018). Whether the same applies to TR in writing is an empirical question.

We should finish by acknowledging a number of limitations of our study. First, we only investigated one task type, and this limits the generalizability of our findings. Second, the observed TR effects (and lack of) need to be taken cautiously given the cross-sectional nature of our study and the lack of WCF provision in the written modality. As for the former, Bygate (2018a, p. 22) has suggested that future research should explore "the degree to which fluency, accuracy, complexity and other aspects of performance might each develop on different iterations". Finally, our limited number of participants limits the potential implications and generalizability of findings.

Despite these limitations, we would like to think that our study adds a relevant piece to studies of task repetition as well as to current debates on how and why writing may be conducive to language learning.

Funding

References

Ahmadian, M. J. (2011). The effect of "massed" task repetitions on complexity, accuracy, and fluency: Does it transfer to a new task? *Language Learning Journal*, 39(3), 269–280. https://doi.org/10.1080/09571736.2010.545239

Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15(1), 35–59. https://doi.org/10.1177/1362168810383329

Agustín-Llach, M. P. (2011). *Lexical errors and accuracy in foreign language writing*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847694188

Amiryousefi, M. (2016). The differential effects of two types of task repetition on the complexity, accuracy, and fluency in computer-mediated L2 written production: A focus on computer anxiety. *Computer Assisted Language Learning*, 29(5), 1052–1068. https://doi.org/10.1080/09588221.2016.1170040

Bardovi-Harling, K. (1992). The relationship of form and meaning: A cross sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics*, 13, 253–278. https://doi.org/10.1017/S0142716400005634

Bardovi-Harling, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in second language acquisition*, 11, 17–34. https://doi.org/10.1017/S0272263100007816

Birjandi, P., & Ahangari, S. (2008). Effects of task repetition on the fluency, complexity and accuracy of Iranian EFL learners. *The Asian EFL Journal*, 10(3), 28–52.

Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.32.02bul

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). Oxford: Macmillan.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 23–48). London: Longman.

Bygate, M. (2006). Areas of research that influence L2 speaking instruction. In E. Usó-Juan & A. Martínez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 159–186). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197778.3.159

Bygate, M. (2018a). Introduction. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 1–25). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.intro

Bygate, M. (Ed.) (2018b). *Learning language through task repetition*. Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11

Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37–74). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.05byg

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Daller, H., van Hout, R., & Trefers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24, 197–222. https://doi.org/10.1093/applin/24.2.197

De Jong, N., & Tillman, P. (2018). Grammatical structures and oral fluency in immediate task repetition. Trigrams across repeated performances. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 43–73). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.02jon

Ellis, R. (2005). Planning and task-based research: Theory and research. In R. Ellis (Ed.). *Planning and task performance in a second language*. (pp. 3–34). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.03ell

Ellis, R. (2012). *Language teaching research and language pedagogy*. Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781118271643

Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford: OUP.

Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–193). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.11ell

Fukuta, J. (2015). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321–340. https://doi.org/10.1177/1362168815570142

Gass, S. M., Mackey, A., Álvarez-Torres, M. J. & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49(4), 549–581. https://doi.org/10.1111/0023-8333.00102

Gilabert, R. (2005). Evaluating the use of multiple sources and methods in needs analysis: A case study of journalists in the autonomous community of Catalonia (Spain). In M. H. Long (Ed.), *Second language needs analysis* (pp. 182–99). Cambridge: CUP. https://doi.org/10.1017/CBO9780511667299.007

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 production. *IRAL*, 45, 215–240. https://doi.org/10.1515/iral.2007.010

Guiraud, P. (1954). *Les Charactères Statistiques du Vocabulaire. Essai de méthodologie*. Paris: Presses Universitaires de France.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. <http://www.victoria.ac.nz/lals/resources/range.aspx> (26 May, 2020).

Hu, X. (2018). Effects of task-type, task-type repetition and performance criteria on L2 oral production. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 149–175). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.06hu

Kim, Y., Crossley, S., Jung Kyle, K. & Kang, S. (2018). The effects of task repetition and task complexity on L2 lexicon use. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 75–97). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.03kim

Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, 41, 829–840. https://doi.org/10.1016/j.system.2013.08.005

Kobayashi, E., & Kobayashi, M. (2018). Second language learning through repeated engagement in a poster presentation task. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 229–260). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.09kob

Kormos, J. (2014). Differences across modalities on performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes & R. M. Manchón (Eds.) *Task-based language learning. Insights from and for L2 writing* (pp. 193–216). Amsterdam: John Benjamins.

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching*, 45, 261–284. https://doi.org/10.1515/iral.2007.012

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Boston, MA: The MIT Press.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15, 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45, 36–62. https://doi.org/10.5054/tq.2011.240859

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, 29, 16–27. https://doi.org/10.1016/j.jslw.2015.06.003

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4, 221–250. https://doi.org/10.1177/136216880000400303

Lynch, T., & MacLean, J. (2001). Effects of the benefits of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching, and testing* (pp. 141–162). London: Longman.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19, 85–104. https://doi.org/10.1191/0265532202lt210a

Manchón, R. M. (2011a). The language learning potential of writing in foreign language contexts: Lessons from research. In T. Cimasko & M. Reichelt (Eds.), *Foreign language writing instruction: Principles and practices* (pp. 44–64). Anderson, SC: Parlor Press.

Manchón, R. M. (2011b). *Learning to write and writing to learn in an additional language*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31

Manchón, R. M. (2014a). The distinctive nature of task repetition in writing. Implications for theory, research, and pedagogy. *ELIA*, 14,13–42. https://doi.org/10.12795/elia.2014.i14.02

Manchón, R. M. (2014b). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights to and from writing* (pp. 27–52). Amsterdam: John Benjamins.

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Matsumara, M., Kawamura, K., & Affricano, A. (2008). Narrative task type repetition and changes in second language use in a classroom environment: A case study. *The Educational Sciences*, 10, 124–145.

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15. https://doi.org/10.1016/j.jslw.2015.06.004

McNamara, D., Graesser, A., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511894664

Mojavezi, A. (2013). The relationship between task repetition and language proficiency. *Applied Research on English Language*, 3(1), 29–40.

Murphy, L., & Roca de Larios, J. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL learners. *Journal of Second Language Writing*, 19, 61–81. https://doi.org/10.1016/j.jslw.2010.02.001

Nitta, R., & Baba, K. (2014). Task repetition and L2 writing development: A longitudinal study from a dynamic systems perspective. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 107–136). Amsterdam: John Benjamins.

Nitta, R., & Baba, K. (2018). Understanding benefits of repetition from a complex dynamic systems perspective. The case of a writing task. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 285–316). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.11nit

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. https://doi.org/10.1093/applin/amp044

Plonsky, L. (2015). Statistical power, *p* values, descriptive statistics and effect sizes: A "back-to-basics" approach to advancing quantitative methods in L2 research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 23–45). New York, NY: Routledge. https://doi.org/10.4324/9781315870908-3

Roca de Larios, J., Manchón, R. M., & Murphy, L. (1996). Generating text in native and foreign language writing: A temporal analysis of problem solving formulation processes. *Modern Language Journal*, 90(1), 100–114. https://doi.org/10.1111/j.1540-4781.2006.00387.x

Sheppard, C., & Ellis, R. M. (2018). The effects of awareness-raising through stimulated recall on the repeated performance of the same task and on a new task of the same type. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 177–198). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.07she

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532. https://doi.org/10.1093/applin/amp047

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394–430. https://doi.org/10.1111/lang.12228

Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Appendix
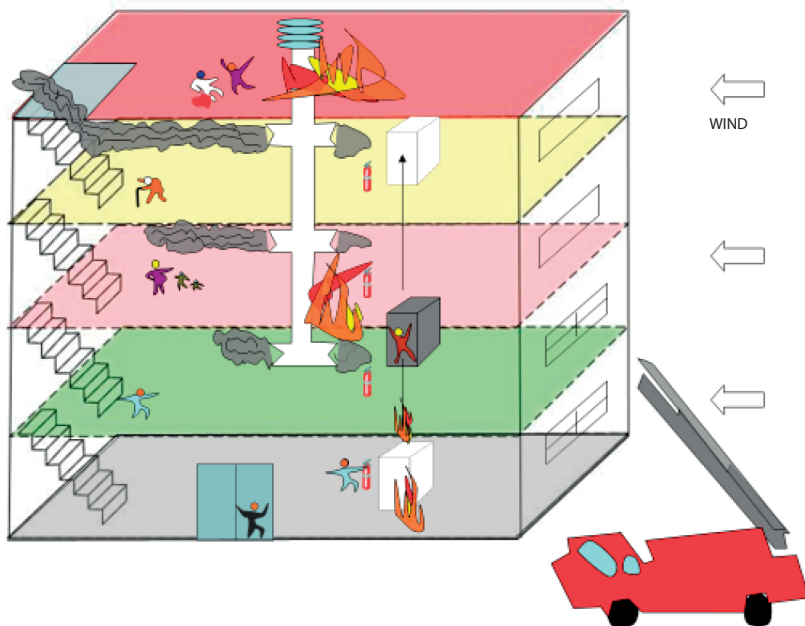
Effects of task repetition with the aid of direct and indirect written corrective feedback

A longitudinal study in an out-of-school context

Victoria Amelohina, Florentina Nicolás-Conesa
and Rosa M. Manchón
University of Murcia

This study sheds light on the learning affordances of task repetition in writing by analyzing the effects of exact task repetition (2 iterations of the same task) and procedural task repetition (2 task types, each one performed twice over 6 months) aided with written corrective feedback (WCF). The participants were 19 EFL students enrolled in an out-of-school instructional program. Two sequences of direct and indirect WCF were implemented and potential differences in terms of CAF measures of the texts produced in a new iteration of the task and across tasks and time were examined. Results showed linear and non-linear effects of task repetition with WCF on diverse components of CAF, as well as a differential appropriation of indirect WCF across time. Theoretical and methodological implications for the learning affordances of task repetition in the environment of writing are drawn.

Introduction

The study to be reported in this chapter intends to add to previous empirical work on the learning potential of task repetition implemented with written corrective feedback (WCF) provision and processing. In what follows we elaborate on the aims and relevance of our study by situating it in the relevant literature.

Task repetition, language learning, and language learning through writing

Task repetition (TR) is a task variable that has attracted theoretical, empirical, and pedagogical attention in the task-based language teaching (TBLT) literature. In the introduction to his recent edited collection on the topic, Bygate (2018) characterizes TR by stating that «'task repetition' does not refer to the precise repetition of the

language used. […] Rather *what is repeated is a given configuration of purposes and a set of content information*» (p. 1. Emphasis added). Along the same lines, Nitta and Baba (2018) contend that "a task should not be considered as a tool to force learners to follow predicted steps but rather be seen as a pedagogical approach to creating an environment in which they can use and learn the L2 in their own autonomous ways" (p. 310). These considerations are at the basis of Larsen-Freeman's (2018) suggestion to replace "task repetition" with "task iteration" in an attempt to avoid "the misleading impression that using the same task more than once will elicit the same response each time" (p. 323). In line with Larsen-Freeman's (2018) suggestion, we shall use the term "iteration" for the second performance of the same task, as detailed in later sections.

The connection between TR and language learning has been theorized mainly in relation to oral tasks although it has also featured more recently in discussions of writing as a site for language learning (e.g. Manchón, 2014 a, b). In the oral domain, TR learning effects are predicated on psycholinguistic and pedagogical grounds. From a psycholinguistic perspective, the learning affordances are premised on the role of working memory and the limitations of attentional resources during real-time, speech production processes: As different task constraints compete for attention when language is produced in real time, having the opportunity to repeat a task is purported to facilitate heightened attention to language during the second iteration because this new opportunity "can help to give learners space to work on matching meanings to language" (Bygate, 2006, p. 172). From a pedagogical perspective, the learning effects of TR have recently been accounted for in terms of regularization of the task environment and, resulting from it, in terms of the associated learner's prediction of the language needed to express one's intended meaning (Bygate, 2018).

The relevance of these pedagogical arguments in support of TR across language modalities is uncontroversial. In contrast, the psycholinguistic rationale for the learning effects of TR rests on the consideration of the criterial characteristics of oral communication, which in part explains why the bulk of extant empirical work has targeted primarily repetition of oral tasks (but see Baba & Nitta, 2012; Nitta & Baba, 2014, 2018, for notable exceptions). Therefore, it is pertinent to question whether or not the same potential learning effects can be posited for the written mode, especially considering two relevant characteristics of writing, namely, "the distinct nature of the temporal dimension of written communication, on the one hand, and the characteristic features of the provision and processing of feedback in the environment of writing, on the other" (Manchón, 2014a, p. 19).

With respect to the temporal dimension – writing takes place off-line, in contrast to the on-line nature of speaking –, the processing constraints that characterize speaking do not apply to writing, perhaps with the exception of various forms of

synchronous computer-assisted communication. Although it is true that writing is a problem-solving communicative event in which task demands are varied and likewise constantly compete for attention, the off-line nature of written communication would in principle allow a more flexible allocation of attentional resources to various task constraints and goals (including the attention to language during the first iteration of the task) as well as greater monitoring during the entire writing process (see Manchón, 2014b). These affordances of writing are also linked to the "internal task repetition" nature of the writing process (Manchón, 2014b), i.e., to the recursive, cyclical (rather than linear) interaction of writing processes (planning, transcribing, evaluating, revising). This means that in writing the task environment is likely to change within the task cycle, which has been documented to result in changes in task representation, in the goals pursued, in the cyclical allocation to attentional resources to diverse task goals throughout the composing process, and, ultimately, in a kind of internal TR that is specific to the written mode.

Additionally, classroom language learners often repeat the same writing task after having received feedback, another key element to be added to the specificity of TR effects in writing, although the role of feedback has in effect been part of TR debates. For instance, Ellis (2009) drew attention to the enhanced learning opportunities that could derive from the combined effect of TR and feedback provision, an argument more recently developed by Bygate (2018), who has suggested that "task repetition may also be important in relation to the feedback that TBLT generates for learners" given that "repeated iteration of tasks could help to provide opportunities for learners to mobilise feedback from previous encounters in the context of upcoming iterations" (p. 12). He further contends that "Ensuring learners can encounter a given task on repeated occasions may create opportunities for progressive internalization of different aspects of the feedback, by enabling attention to be cyclically focused and re-focused" (Bygate, 2018, p. 12). How and why this cyclical appropriation of feedback takes place over time is an empirical question in any modality, and certainly in the case of writing, a domain in which the research priority in feedback studies has been the study of its effects in immediate revisions (as reviewed in Bitchener & Storch, 2016).

A pedagogically-relevant concern would therefore be to investigate the predicted/observed effects of TR implemented as an intervention in the form of feedback when writing, receiving feedback, and rewriting is fully embedded in the students' learning experience (see Chapters 14 and 17), as we did in our own research (see below).

Methodological considerations

In our study we investigated the longitudinal appropriation and effects of two forms of written corrective feedback (WCF) in both exact task repetition (2 iterations of the same tasks) and procedural task repetition (2 task types, each one performed twice during a 6 month instructional program). Our decision to investigate learners' improvement in task performance in terms of two forms of WCF took account of the conflicting findings on the effects of direct and indirect WCF on immediate revisions (see review in Bitchener & Storch, 2016). Our study sheds new light on the appropriation and effects of two forms of WCF within a single task (as done in most extant WCF research) and across tasks and time (the suggested research path to be followed). In other words, we addressed TR affordances in terms of repetition of the same task and task-type repetition, and did so longitudinally

Regarding the longitudinal perspective adopted, Bygate (2018) suggests that relevant questions for future research include investigating which dimensions of task performance change across iterations and tasks, and whether any observed changes in any of the CAF dimensions happen in parallel. Adding to this, and specifically in the area of writing, Nitta and Baba (2014) have argued that TR affordances may come about as a result of repeated use of tasks over time, hence their claim that "it is vital to have a sufficient number of data points to capture the process of changes" (Baba & Nitta, 2014, p. 28).

The potential task-dependency of TR effects was another variable in our design as TR effects in the area of writing have been shown to be mediated by task type (Nitta & Baba, 2014, 2018). However, given the lack of empirical evidence of the carryover of effects of TR to a new (similar or identical) task, open questions exist as to whether the key variable is task repetition *per se* or task-type repetition. In the only study investigating this issue in writing, Nitta and Baba (2014) concluded that whereas the effects of repeating the same task may be minimal, task-type repetition does result in improvements in the language used at the levels of lexis and grammar. We intended to test this prediction and to do so when the massed repetition practice mentioned above was built into the instructional program.

The present study: Research questions

Building on previous TR research, we designed a longitudinal study in which we investigated the effects of TR aided with two forms of WCF (direct and indirect) on four reading-to-write tasks (2 task types, each one repeated twice) on the characteristics of the texts produced in terms of CAF measures. This overall aim was operationalized in terms of three research questions:

1. Does TR with the help of direct and indirect WCF result in differences in the accuracy of the text produced, in terms of both global accuracy measures and specific error types? Do any observed effects vary across the same task iterations and across tasks and time?
2. Does TR with the help of direct and indirect WCF result in differences in fluency? Do any observed effects vary across the same task iterations and across tasks and time?
3. Does TR with the help of direct and indirect WCF result in differences in lexical and syntactic complexity? Do any observed effects vary across the same task iterations and across tasks and time?

Method

Participants and context

The participants were 19 students, 7 males and 12 females, aged between 22 and 40, who had been studying English for approximately 12 years. They were enrolled in an intermediate level, 6-month course at a private language school in a Spanish city with the purpose of taking the Trinity exam of ISE II (B2 level according to the Common European reference Framework) at the end of the course. The stakes for passing this exam in a timely manner were very high for the participants as they were employed primary school teachers who were required by their school administration to obtain the B2 level in English to keep their jobs.

They were divided into 4 groups taught by the same teacher (the first researcher). During the 24-week (48 hours) course, the four language skills and grammar contents of the B2 level were covered. Participants had two contact hours per week, which were devoted to the improvement of their English skills and exam preparation. In addition, participants also had to complete homework activities to consolidate what they have learnt in class. Due to ethical reasons, there was not a control group since all the participants had to pay a monthly fee for the course. This decision nevertheless adds to the ecological validity of the study at the expense of constituting a potential threat to its internal validity. The participants accepted to take part in the study on a voluntary basis and they all signed an informed consent form.

Data collection procedures

Data were collected at four different points during the 6 months duration of the course. In addition to their regular class work and assignments, and outside their regular contact hours, the participants completed four 60-minute, tasks

(approximately 250 words) without using external sources such as dictionaries or the Internet, hence reproducing the Trinity exam conditions. The students were asked to write an **article** at Time 1 (pollution and recycling) and an **article** again at Time 3 (home schooling). At Time 2 and 4 they wrote 2 different **reports** about the world of work. In order to accomplish these tasks, they had to read a text that was given to them as a written prompt and subsequently write an article (Task 1 and Task 3) and a report (Task 2 and Task 4), summarising the main ideas of the written prompt and developing the texts based on the information provided in the source material. These task types and topics were part of their curriculum and exam preparation requirements. Two days after having written their texts, the participants received comprehensive WCF (direct or indirect) on their language errors. Direct written corrective feedback (DWCF) consisted in crossing out the error and providing the correct L2 form, while in the case of indirect written corrective feedback (IWCF) the errors were highlighted using different codes for the various categories (grammar, lexis, spelling, or punctuation.

Table 1. Data collection

Tasks	Writings	Written assignment	Feedback received	
			Group 1 (Sequence 1)	Group 2 (Sequence 2)
Task 1 (Weeks 7–8)		Article (pollution and recycling)	DWCF	IWCF
Task 2 (Weeks 15–16)		Report (the world of work)	IWCF	DWCF
Task 33 (Weeks 19–20)		Article (home schooling)	DWCF	IWCF
Task 4 (Weeks 23–24)		Report (the world of work)	IWCF	DWCF

We created sequences of feedback and the provision of DWCF and IWCF was counterbalanced across tasks for all participants, as shown in Table 1. Feedback provision was followed by a feedback processing stage, which was completed at home. The participants were asked to identify and write their errors in a table and to attempt to explain the reason for their errors. In case they did not know how to solve the problems and errors identified, they were encouraged to discuss it with their teacher. For this purpose, 30 minute individual tutorials were organised for students. The WCF processing data are not analysed in the present study.

Five days after having received and processed the WCF, the participants were asked to repeat the original task in the same conditions as during the first performance.

Data analysis

We analyzed the complexity accuracy, and fluency (CAF) of our participants' texts in the two iterations of the four writing tasks. Although the WCF provided targeted only accuracy, we decided to look into all CAF dimensions for two reasons. One relates to the divergent findings on the effects of TR on different dimensions of performance reported in previous TR research (in both speaking and writing). We also targeted all CAF measures in an attempt to add to previous work on TR in writing that has traced changes over time, which is rather limited (cf. Baba & Nitta, 2014; Nitta & Baba, 2014, 2018) and to address the request "to include accuracy together with fluency and complexity measures to gain a better understanding of dynamic L2 development" (Nitta & Baba, 2014, pp. 127–128) in studies of TR in writing.

For complexity, we used the Synlex program. Linguistic complexity involved the distinction between lexical and syntactic complexity as "separate, independent dimensions of L2 performance and L2 proficiency" (Bulté & Housen, 2014, p. 53). Within lexical complexity, lexical density (LD: the ratio of lexical words), and lexical sophistication (LS: the proportion of advanced words) were examined. Syntactic complexity included coordination, subordination, nominal complexity, as well as diversity of syntactic structures (Norris & Ortega, 2009). Aiming at covering all dimensions of production, but avoiding redundancy (Norris & Ortega, 2009; Bulté & Housen, 2012, 2014), we measured syntactic complexity via coordination (dividing the number of coordinate phrases by the number of clauses: CP/C) and via subordination (dividing the number of dependent clauses by the number of clauses: DC/C). We also computed a general measure of syntactic complexity (mean length of T-unit: MLT), as well as nominal complexity, as measured through the mean length of clause (MLC) and complex nominals per clause (CN/C).

Accuracy was analyzed with the help of a coding system for error types previously developed by our research team (see Nicolás-Conesa, Manchón, & Cerezo, 2019). The coding system distinguished three superordinate categories (broad coding), following Van Beuningen, De Jong, & Kuiken (2012), and Ferris, Liu, Sinha, & Senna (2013): (a) *grammatical errors*: (b) *lexical* errors, and (c) *orthographical errors*. The categories included within each of these superordinate categories were considered part of a narrow coding (e.g. word form, word order, word choice, spelling, etc).

The first two researchers analysed 30% of the data with high inter-rater reliability (broad coding: .95 Cronbach's Alpha; narrow coding: .92 Cronbach's Alpha). The first author subsequently continued with the full coding process on her own. The high inter-rater reliability obtained can be explained by the two coders' previous familiarity with the coding schemes.

We computed the global percentage of error rate for 2 iterations in each of the 4 tasks in terms of total number of errors/total number of words *100, which

was considered an objective measure of analysis. We did not use T-units because their validity to measure linguistic accuracy has been questioned (Bardovi-Harlig & Bofman, 1989) since T-units do not control for the number of errors included in each T-unit (i.e. one error or multiple errors) or the length of T-units. For the analysis of broad error types (grammatical, lexical, and orthographical), we used the raw number of errors in each iteration of the four tasks. As regards fluency, we computed the total number of words and sentences written in each iteration of the tasks.

We conducted mixed between- and within- subjects ANOVAs to compare the impact of two sequences of WCF on our dependent variables (complexity, accuracy, and fluency) across four tasks and two iterations of the same task. We used for our ANOVAs one categorical independent between-subjects variable (sequence of feedback, with 2 levels: Sequence 1 and Sequence 2) and two categorical independent within-subjects variables (task: 4 levels – tasks 1 to 4-; writings (iterations): 2 levels – writing 1 and writing 2 of each of the 4 tasks).

Results

Our first research questions asked whether TR with the help of direct WCF (DWCF) and indirect WCF (IWCF) resulted in differences accuracy (in terms of global accuracy measures and specific error types), as well as whether any potential effects varied across task iterations (referred to as "Writings" in what follows) and across the four tasks. The results will be reported bearing in mind that the two groups received two different sequences of feedback, namely, Sequence 1: Task 1: DWCF; Task 2: IWCF; Task 3: DWCF; Task 4: IWCF; Sequence 2: Task 1: IWCF; Task 2: DWCF; Task 3: IWCF; Task 4: DWCF. We will display the significant effects in tables. For clarification purposes, we will also show non-significant effects (such as no effects for different groups or sequences) when relevant.

Regarding global accuracy measures, we found significant differences in the global percentage of errors across the four tasks. Both WCF sequences reduced their errors across Tasks (Tables 2 and 3; Figures 1 and 2) and there were no significant differences between groups (Task*Sequence) in this decrease of errors.

Table 2. ANOVA results for global percentage of errors across tasks and writings for both sequences

Effects	Wilks lambda	F	Hypothesis df	Error df	p	Effect size $\eta^2 p$
Task	.52	4.67	3	15	.02	.48
Task * Sequence	.86	.85	3	15	.49	.15
Writings	.49	17.13	1	17	.00	.50
Writings * Sequence	.98	.42	1	17	.52	.02

Table 3. Global percentage of errors across tasks and writings for both Sequences

| | TASK 1 | | | TASK 2 | | | TASK 3 | | | TASK 4 | | |
| | Wr1 | Wr2 | Wr1 & Wr2 (Task 1) | Wr1 | Wr2 | Wr1 & Wr2 (Task 2) | Wr1 | Wr2 | Wr1 & Wr2 (Task 3) | Wr1 | Wr2 | Wr1 & Wr2 (Task 4) |
	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
1	22.11 (10.72)	14.71 (7.22)	18.41 (7.94)	15.82 (0)	16.35 (8.13)	16.08 (7.78)	19.88 (8.99)	11.57 (7.35)	15.73 (7.03)	14.73 (8.03)	11.44 (6.33)	13.08 (6.71)
2	18.40 (10.08)	15.14 (8.82)	16.77 (9.06)	19.27 (9.70)	12.66 (5.84)	15.97 (7.23)	16.03 (6.12)	13.77 (7.36)	14.90 (6.58)	14.83 (6.58)	13.51 (5.78)	14.17 (6.08)

NOTE. Wr1: writing 1; Wr2: writing 2
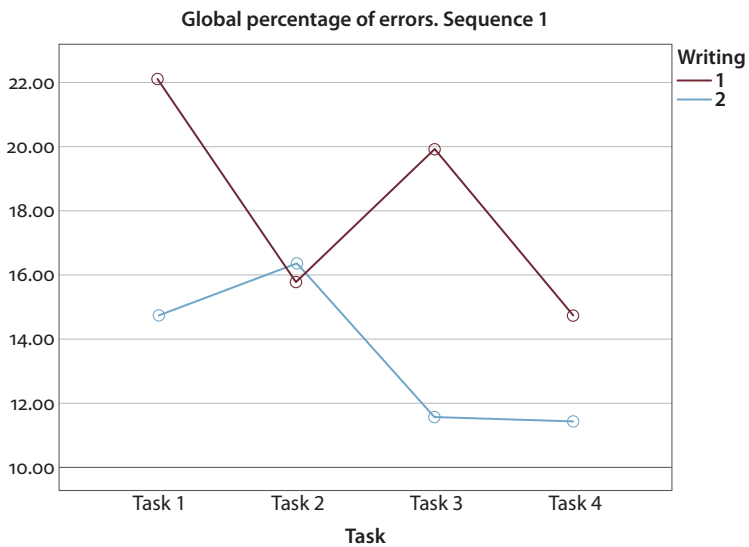
Figure 1. Global percentage of errors across tasks and writings. Sequence 1
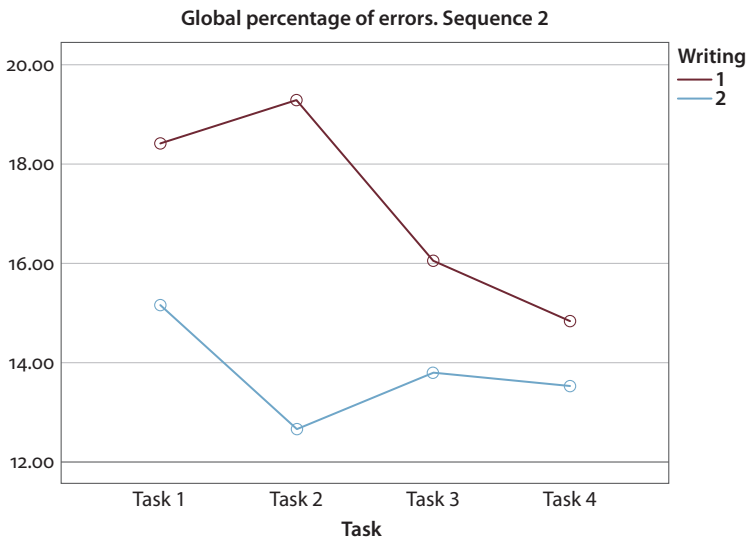
Figure 2. Global percentage of errors across tasks and writings. Sequence 2

There was also a significant decrease of errors across Writings (task iterations) for both sequences (see the blue lines corresponding to task iterations in Figures 1 and 2). No significant differences were found between groups in the decrease of errors across Writings (iterations) (writings*sequence) (see Tables 2 and 3).

Our first research question also asked about the effects of TR across Writings (task iterations) and across tasks and time in terms of specific error types. Regarding grammar errors, we found a significant effect for Task in both sequences (Table 4): Grammar errors followed a similar trajectory in both sequences. The mean of grammar errors increased across Task 2 and Task 3, and in Task 4 grammar errors decreased to a slightly lower level than at Task 1 (see Table 5 and Figures 3 and 4).

There was also an interaction effect among Task* Writing* Sequence regarding grammar errors (Table 4). In both sequences of feedback there was an immediate increase of errors in the second iteration of the task after having received IWCF for the first time, which corresponded to the second performance of Task 2 for Sequence 1, and to the second performance of Task 1 for Sequence 2 (see Table 5 and the blue lines in Figures 3 and 4). In contrast, the second time that both sequences received IWCF there was a decrease of errors in the second iteration of the task.

Table 4. ANOVA results for grammar and non-grammar errors across tasks and writings for both sequences

Variables	Effects	Wilks lambda	F	Hypothesis df	Error df	p	Effect size $\eta^2 p$
Grammar errors	Task	.52	4.56	3	15	.02	.48
	Task*Writings*Sequence	.49	5.05	3	15	.01	.50
Non-grammar errors	Task	.49	5.02	3	15	.01	.50
	Writings	.78	4.83	1	17	.04	.22

Table 5. Grammar and non-grammar errors across tasks and writings for both sequences

Variables	Sequence	TASK 1			TASK 2			TASK 3			TASK 4		
		Wr1	Wr2	Wr1 & Wr2 (Task 1)	Wr1	Wr2	Wr1 & Wr2 (Task 2)	Wr1	Wr2	Wr1 & Wr2 (Task 3)	Wr1	Wr2	Wr1 & Wr2 (Task 4)
		M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
Grammar errors	1	21 (9.29)	18.80 (10.67)	19.90 (7.85)	17.10 (9.07)	24.00 (11.21)	20.55 (8.66)	32.40 (13.13)	17.10 (10.34)	24.75 (8.84)	20 (10.96)	18 (11.83)	19 (10.26)
	2	19.78 (9.36)	22.78 (10.17)	21.28 (8.73)	26.89 (14.41)	18.56 (8.62)	22.72 (10.58)	25.78 (7.40)	23.67 (11.57)	24.72 (8.72)	20.33 (9.04)	19.11 (7.79)	19.72 (7.87)
Non-grammar errors	1	11 (6.99)	11 (4.74)	11 (4.58)	15.10 (8.86)	16.30 (9.97)	15.70 (9.04)	17.20 (11.59)	11.60 (8.87)	14.40 (9.50)	15.80 (11.83)	10.30 (5.56)	13.05 (8.18)
	2	11.11 (8.48)	8.22 (6.06)	9.67 (6.62)	15.44 (9.95)	11.78 (5.49)	13.61 (6.74)	12.78 (8.53)	11.89 (7.74)	12.33 (7.95)	14.33 (9.42)	13.22 (9.11)	13.78 (7.74)
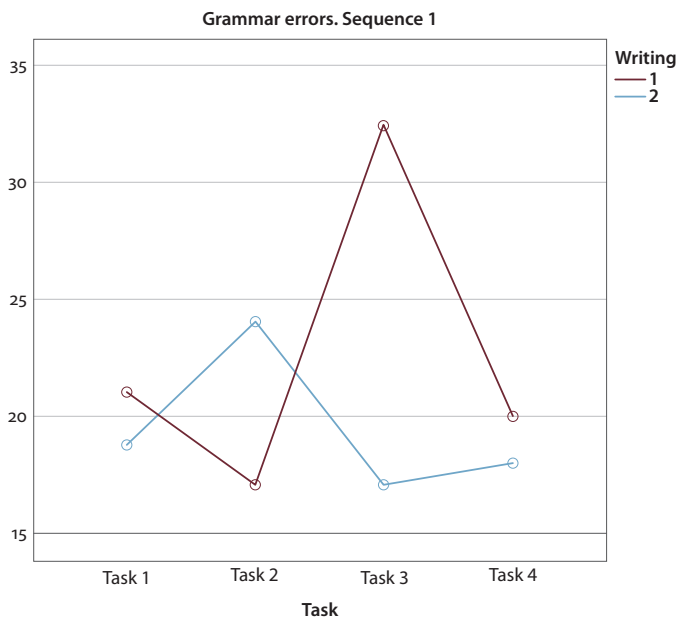
Figure 3. Grammar errors across tasks and writings. Sequence 1
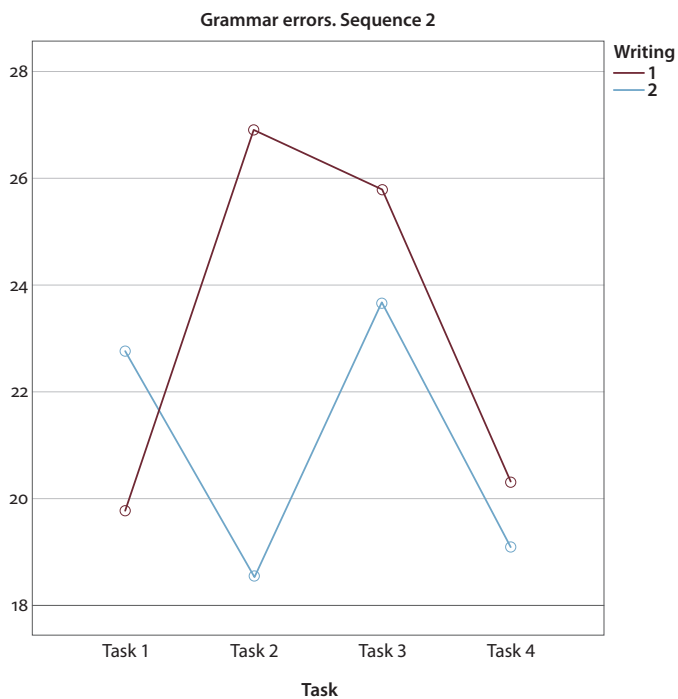
Figure 4. Grammar errors across tasks and writings. Sequence 2

As for non-grammar errors, there was a main effect for Task in both sequences of feedback (Table 5). We found an increase of non-grammar errors from Task 1 to Task 2, and from Task 1 to Task 4 for both sequences (see Table 5 and Figures 5 and 6). However, the number of non-grammar errors decreased across Writings (task iterations) for both sequences of feedback (see Table 5 and Figures 5 and 6).
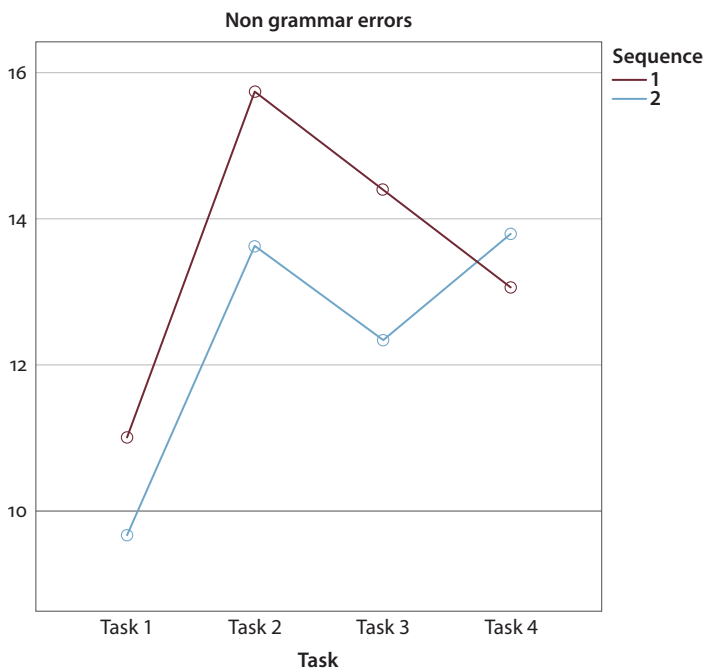
Figure 5. Non-grammar errors across tasks for both sequences

Within non-grammar errors, we further distinguished between lexical errors and other errors, which were merged into one single category/variable (spelling and punctuation). There were no significant differences for lexical errors across Tasks or across Writings (task iterations) (Table 6) for any of the two WCF sequences. This finding could be explained by the low frequency of lexical errors in Task 1 for both sequences.

      We also found an interaction effect between Writings*Sequence that indicated different tendencies between sequences (Table 6). Sequence 1 tended to decrease lexical errors across task iterations, probably because the participants in this sequence started with higher number of lexical errors, and the participants in Sequence 2 tended to slightly increase their lexical errors across task iterations (see the mean number of lexical errors in the iterations of Task 3 and Task 4, Table 7).
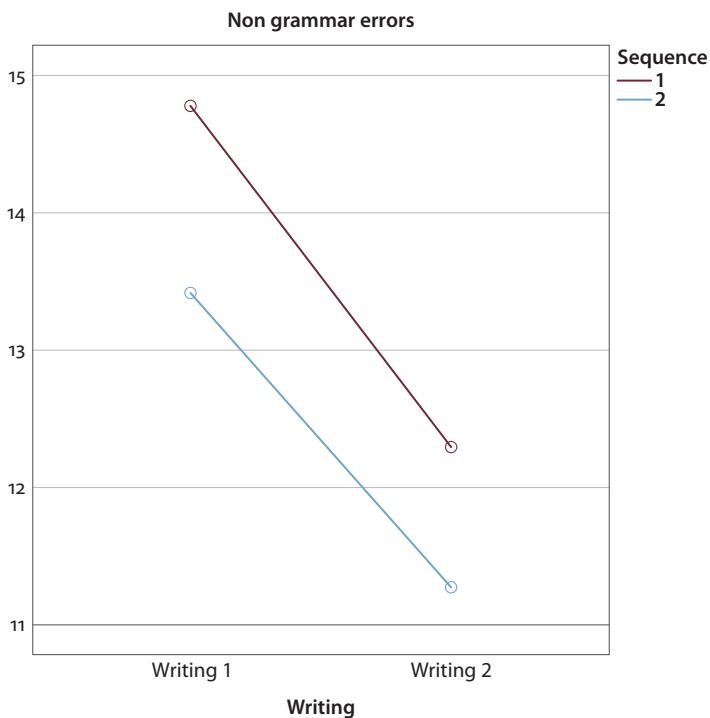
Figure 6. Non-grammar errors across writings for both sequences

Table 6. ANOVA results for lexical and spelling+punctuation errors across tasks and writings for both sequences

Variables	Effects	Wilks lambda	F	Hypothesis df	Error df	p	Effect size $\eta^2 p$
Lexical errors	Tasks	.63	2.97	3	15	.07	.37
	Writings	.90	1.86	1	17	.19	.09
	Writings*Sequence	.77	5.03	1	17	.04	.23
Spelling+Punctuation errors	Tasks	.52	4.71	3	15	.02	.49
	Writings	.71	6.92	1	17	.02	.29

Table 7. Lexical errors and spelling+punctuation errors across tasks
and writings for both sequences

Variables	Sequence	TASK 1			TASK 2			TASK 3			TASK 4		
		Wr1	Wr2	Wr1 & Wr2 (Task 1)	Wr1	Wr2	Wr1 & Wr2 (Task 2)	Wr1	Wr2	Wr1 & Wr2 (Task 3)	Wr1	Wr2	Wr1 & Wr2 (Task 4)
		M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
Lexical errors	1	4.10 (2.38)	4.80 (2.82)	4.45 (1.83)	5.20 (1.81)	5.90 (2.92)	5.55 (2.13)	7.50 (5.08)	3.70 (2.91)	5.60 (3.23)	6.70 (3.65)	4.10 (2.69)	5.40 (2.69)
	2	3.67 (1.73)	3.44 (3.36)	3.56 (1.74)	4.78 (4.41)	3.56 (2.65)	4.17 (2.98)	3.33 (2.18)	4.33 (3.20)	3.83 (2.61)	5.00 (2.24)	6.67 (5.50)	5.83 (3.38)
Spelling+Punctuation errors	1	6.70 (6.52)	6.10 (3.18)	6.40 (4.29)	9.60 (7.58)	9.50 (8.06)	9.55 (7.51)	9.50 (7.66)	7.80 (7.15)	8.65 (6.98)	9.10 (10.49)	6.00 (4.16)	7.55 (6.93)
	2	7.44 (8.29)	4.78 (3.53)	6.11 (5.45)	9.78 (6.38)	7.78 (4.97)	8.78 (5.12)	9.00 (6.59)	6.67 (5.52)	7.83 (5.88)	9.11 (7.77)	6.11 (4.65)	7.61 (5.87)

As for the combined variable of spelling and punctuation, we found a significant effect for Task and for Writing in both WCF sequences (Table 6). Both sequences increased their spelling and punctuation errors across tasks, but there was a decrease across task iterations (Table 7).

Our second research question asked about TR effects in the area of fluency. As shown in Tables 8 and 9, both sequences of feedback increased the number of words across Tasks and across iterations (Writings), hence the observed interaction effect (Table 8) between Task*Writing.

Table 8. ANOVA results for words and sentences across tasks and writings for both sequences

Variables	Effects	Wilks lambda	F	Hypothesis df	Error df	p	Effect size $\eta^2 p$
Words	Tasks	.18	22.91	3	15	.00	.82
	Writings	.45	21.01	1	17	.00	.55
	Tasks*Writings	.55	4.14	3	15	.03	.45
Sentences	Tasks*Writings*Sequence	.34	9.64	3	15	.00	.66

We also measured fluency in terms of the number of sentences written. An interaction effect among Tasks*Writings*Sequence was observed (Table 8). There was a tendency for Sequence 1 to increase the number of sentences across tasks and writings, except for task 4, in which the number of sentences decreased in the second performance of the task (see blue line in Figure 7). For Sequence 2, there was also a slight tendency to increase the number of sentences across tasks (see blue line in Figure 8), but the number of sentences decreased in the second performance of tasks 2 and 3 (blue line in Figure 8). As a whole, the number of sentences increased across tasks for both sequences, but the increase was not so large for Sequence 2 (see Table 9 and blue lines in Figures 7 and 8).

Table 9. Words and sentences across tasks and across writings for both sequences

Variables	Sequence	TASK 1			TASK 2			TASK 3			TASK 4		
		Wr1	Wr2	Wr1 & Wr2 (Task 1)	Wr1	Wr2	Wr1 & Wr2 (Task 2)	Wr1	Wr2	Wr1 & Wr2 (Task 3)	Wr1	Wr2	Wr1 & Wr2 (Task 4)
		M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
Words	1	155.70 (46.23)	209.80 (35.95)	182.75 (29.53)	205.60 (30.42)	242.30 (28.89)	223.95 (26.26)	250.30 (15.81)	248.90 (17.99)	249.60 (13.84)	243.20 (17.98)	245.50 (16.93)	244.35 (16.76)
	2	181.44 (48.79)	219.22 (64.36)	200.33 (46.29)	243.67 (105.55)	249.44 (57.46)	246.56 (74.83)	251.22 (24.25)	260.11 (16.95)	255.67 (18.37)	222.22 (30.43)	238.89 (26.47)	230.56 (26.74)
Sentences	1	10.40 (3.57)	12.70 (3.34)	11.55 (3.02)	11 (2.67)	14.40 (2.17)	12.70 (1.98)	15.50 (2.92)	15.30 (3.40)	15.40 (3.09)	14.50 (3.41)	13.50 (2.79)	14 (2.94)
	2	11 (3.35)	12.44 (3.75)	11.72 (3)	14 (4.09)	13.89 (4.19)	13.94 (2.76)	15 (2.83)	13.89 (2.80)	14.44 (2.44)	11.67 (1.87)	13.33 (1.23)	12.50 (1.17)
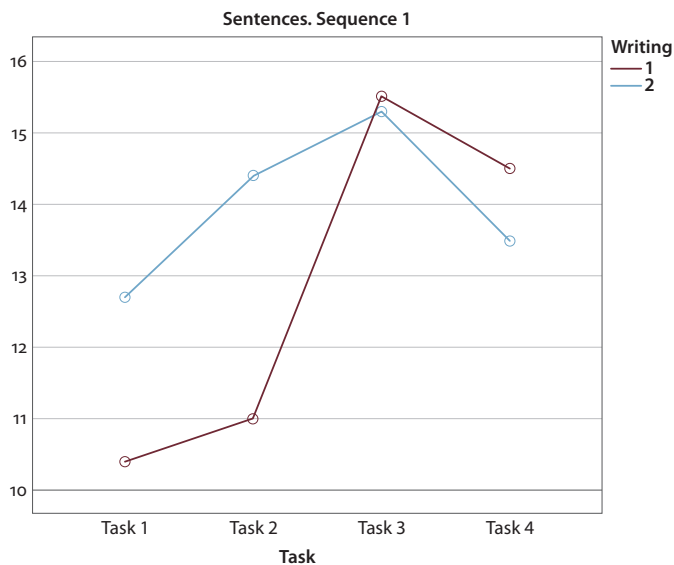
Figure 7. Sentences across tasks and writings. Sequence 1
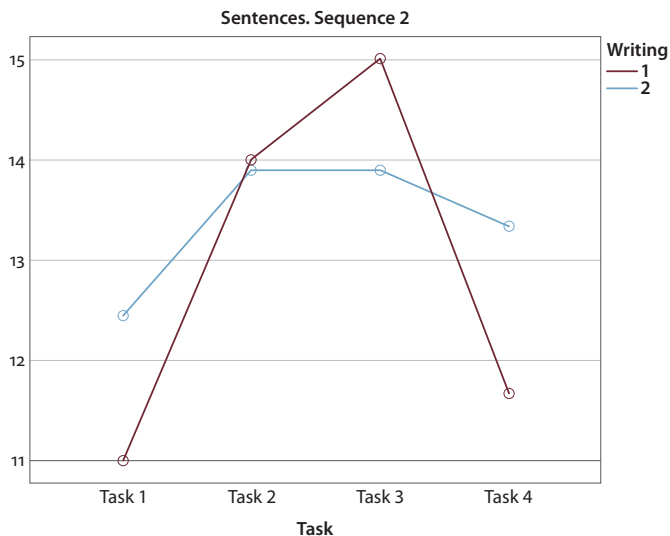
Figure 8. Sentences across tasks and writings. Sequence 2

Our third research question asked about TR effects in the areas of lexical and syntactic complexity. As for lexical complexity, Lexical Density (LD) increased for both sequences across tasks (Tables 10 and 11; Figures 9 and 10)
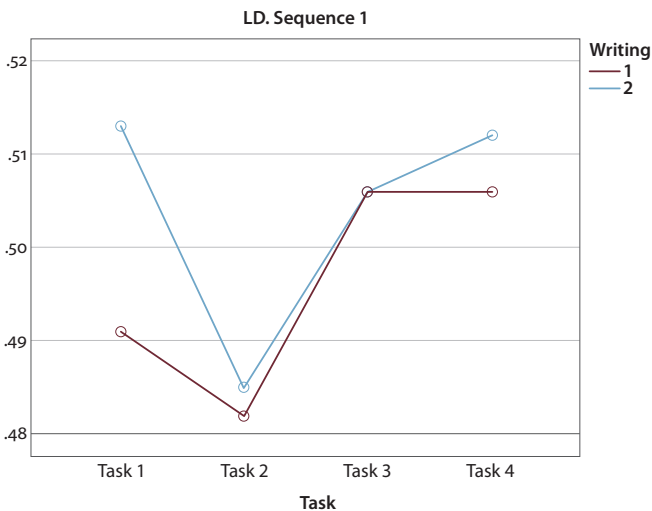
Figure 9. LD across tasks and writings. Sequence 1
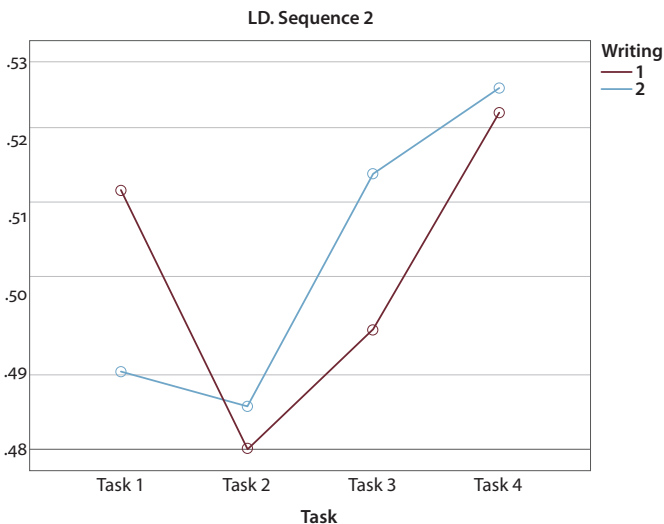
Figure 10. LD across tasks and writings. Sequence 2

It should also be noted that LD decreased for both sequences in Task 2, which could be related to the different types of texts written across tasks (Task 1: article; Task 2: report), although this effect disappeared across time (see the evolution of the blue line across tasks in Figures 9 and 10). On the other hand, lexical sophistication (LS) decreased for both sequences across Tasks (Tables 10 and 11; Figures 11 and 12).

Table 10. ANOVA results for LD and LS across tasks and writings for both sequences

Variables	Effects	Wilks lambda	F	Hypothesis df	Error df	p	Effect size $\eta^2 p$
LD	Tasks	.34	9.83	3	15	.00	.66
LS	Tasks	.25	15.23	3	15	.00	.75

Table 11. LD and LS across tasks and across writings for both sequences

Variables	Sequence	TASK 1			TASK 2			TASK 3			TASK 4		
		Wr1	Wr2	Wr1 & Wr2 (Task 1)	Wr1	Wr2	Wr1 & Wr2 (Task 2)	Wr1	Wr2	Wr1 & Wr2 (Task 3)	Wr1	Wr2	Wr1 & Wr2 (Task 4)
		M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
LD	1	.49 (.02)	.51 (.04)	.50 (.02)	.48 (.02)	.49 (.03)	.48 (.02)	.51 (.02)	.51 (.03)	.51 (.02)	.51 (.04)	.51 (.02)	.51 (.03)
	2	.51 (.02)	.49 (.03)	.50 (.02)	.48 (.04)	.49 (.03)	.48 (.03)	.49 (.03)	.52 (.04)	.51 (.03)	.52 (.04)	.53 (.03)	.53 (.04)
LS	1	.31 (.08)	.31 (.06)	.31 (.06)	.22 (.08)	.23 (.06)	.22 (.06)	.21 (.04)	.19 (.03)	.20 (.03)	.23 (.06)	.22 (.06)	.22 (.05)
	2	.29 (.08)	.28 (.02)	.28 (.05)	.23 (.05)	.19 (.04)	.21 (.04)	.20 (.03)	.17 (.04)	.19 (.03)	.26 (.05)	.24 (.05)	.25 (.04)
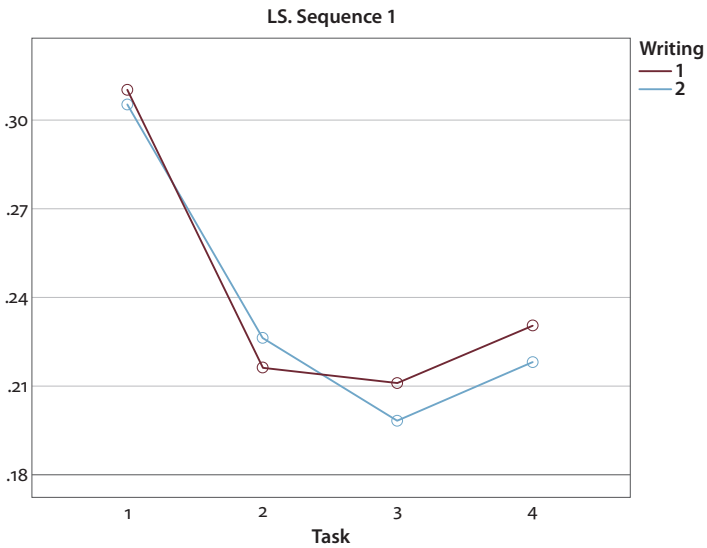
Figure 11. LS across tasks and writings. Sequence 1
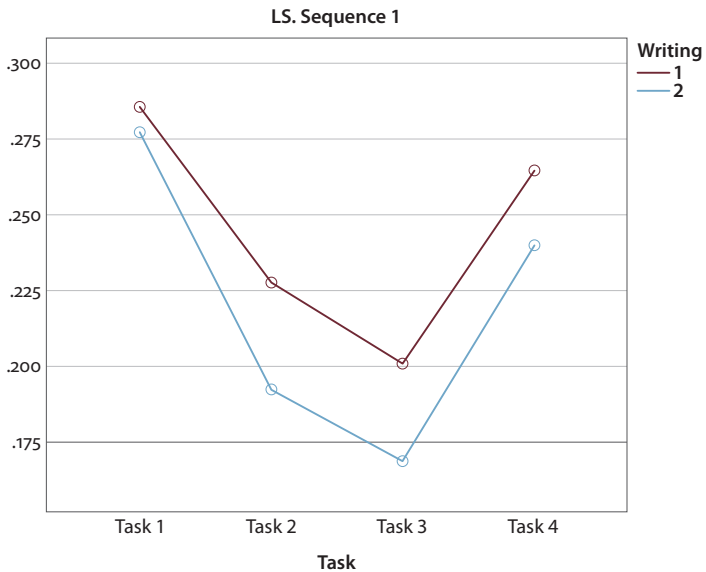
Figure 12. LS across tasks and writings. Sequence 2

With respect to syntactic complexity (Table 12), no significant differences were found in complexity via coordination (CP/C) across Tasks or across Writings (iterations) for any of the two WCF sequences. As for complexity via subordination (DC/C), there was an interaction effect between Task*Writing*Sequence.

Table 12. ANOVA results for syntactic complexity across tasks and writings for both sequences

Variables	Effects	Wilks lambda	F	Hypothesis df	Error df	*p*	Effect size $\eta^2 p$
CP/C	Tasks	.78	1.35	3	15	.29	.21
	Writings	.95	.89	1	17	.36	.05
DC/C	Task*Writings*Sequence	.59	3.49	3	15	.04	.41
MLT	Tasks	.9	.56	3	15	.65	.1
	Writings	.95	.86	1	17	.37	.05
	Task*writings*Sequence	.55	4.09	3	15	.03	.45
MLC	Tasks	.48	5.37	3	15	.01	.52
CN/C	Tasks	.43	6.53	3	15	.01	.57

For Sequence 1, there was a tendency to slightly increase DC/C across tasks, but there were different moves across iterations (see Table 13 and Figure 13). In contrast, for Sequence 2, there was a tendency to increase DC/C across writings (Figure 14) but the ratio of DC/C remained the same across tasks (see Table 13). In spite of the different moves of DC/C across tasks and writings for both sequences, the overall ratio of DC/C remained basically the same across tasks for both WCF sequences.

Table 13. Syntactic complexity across tasks and across writings for both sequences

Variables	Sequence	TASK 1			TASK 2			TASK 3			TASK 4		
		Wr1	Wr2	Wr1 & Wr2 (Task 1)	Wr1	Wr2	Wr1 & Wr2 (Task 2)	Wr1	Wr2	Wr1 & Wr2 (Task 3)	Wr1	Wr2	Wr1 & Wr2 (Task 4)
		M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
CP/C	1	1.36 (3.91)	.16 (.11)	.76 (1.98)	.17 (.08)	.16 (.08)	.16 (.07)	.14 (.07)	.14 (.09)	.14 (.07)	.14 (.08)	.14 (.06)	.14 (.07)
	2	.14 (.13)	.18 (.05)	.16 (.08)	.18 (.08)	.17 (.09)	.17 (.07)	.13 (.09)	.11 (.07)	.12 (.06)	.15 (.09)	.18 (.09)	.16 (.06)
DC/C	1	.35 (.07)	.32 (.08)	.34 (.05)	.42 (.15)	.34 (.12)	.38 (.13)	.36 (.10)	.37 (.09)	.37 (.07)	.34 (.08)	.36 (.11)	.35 (.09)
	2	.35 (13)	.43 (.07)	.39 (10)	.35 (.08)	.37 (.10)	.36 (.07)	.37 (.07)	.39 (.09)	.38 (.08)	.39 (.08)	.36 (.08)	.38 (.08)
MLT	1	133.21 (23.36)	142.17 (16.08)	137.69 (18.13)	166.28 (56.96)	141.75 (40.76)	154.02 (44.69)	145.60 (30.78)	153.78 (31.39)	149.69 (25.72)	151.53 (32.84)	161.12 (40.12)	156.32 (35.89)
	2	158.05 (39.24)	156.12 (33.59)	157.09 (34.69)	143.49 (29.32)	151.38 (42.06)	147.44 (32.77)	145.39 (18.57)	163.85 (39.31)	154.62 (25.14)	163.31 (33.42)	161.59 (26.26)	162.45 (27.88)
MLC	1	84.91 (15.83)	87.50 (11.31)	86.21 (11.22)	83.09 (11.27)	82.94 (6.74)	83.32 (6.99)	88.51 (12.15)	88.64 (13.83)	88.58 (12.59)	96.08 (13.03)	96.08 (13.03)	96.08 (12.29)
	2	86.53 (18.33)	83.50 (13.43)	85.02 (15.67)	87.94 (13.54)	86.89 (11.64)	87.41 (11.82)	88.13 (6.58)	89.05 (17.68)	88.59 (11.09)	94.37 (15.99)	98.07 (11.44)	96.22 (9.31)
CN/C	1	3.74 (4.87)	4.01 (5.18)	3.87 (3.86)	8.15 (5.06)	3.97 (5.04)	6.06 (3.55)	7.26 (5.62)	5.30 (5.77)	6.28 (5.23)	10.53 (5.25)	9.80 (6.53)	10.17 (5.68)
	2	4.77 (6.02)	1.93 (3.55)	3.35 (4.27)	4.81 (6.14)	4.58 (5.76)	4.69 (5.42)	6.66 (5.47)	6.62 (5.62)	6.64 (4.93)	8.59 (6.23)	11.11 (4.31)	9.85 (3.79)
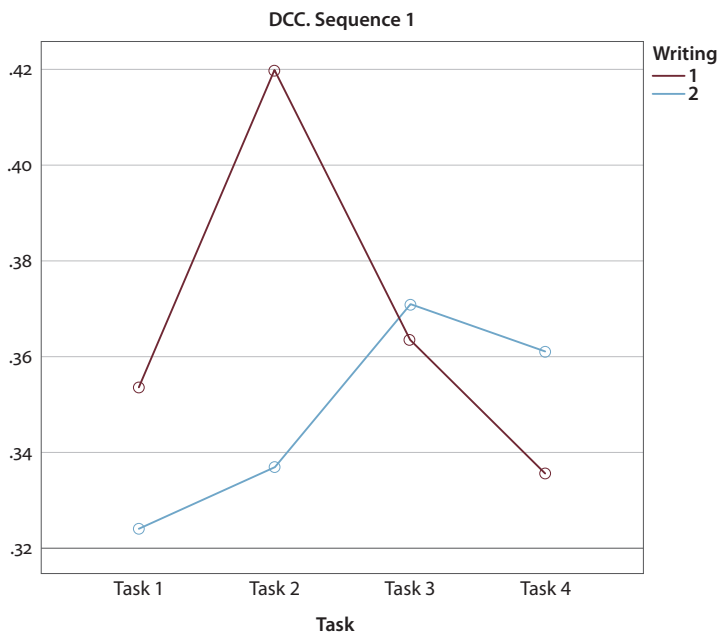
DCC. Sequence 1

Figure 13. DC/C across tasks and writings. Sequence 1
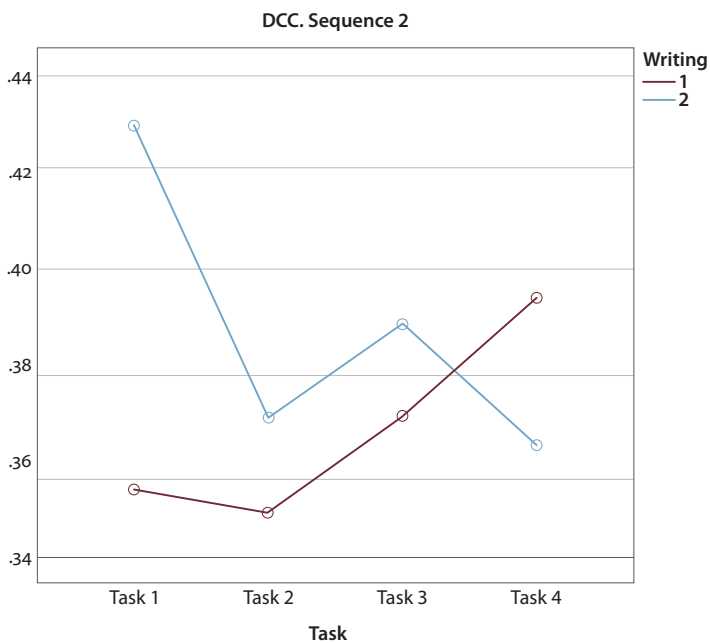
DCC. Sequence 2

Figure 14. DC/C across tasks and writings. Sequence 2

As for the Mean Length of T-unit (MLT), there was an interaction effect between Task*Writing*Sequence (Table 12) Both WCF sequences tended to increase MLT across tasks and writings although the participants in both sequences decreased MLT in iterations of specific (and not always identical) tasks (see Table 13; Figures 15 and 16). Specifically, Sequence 1 decreased MLT in the second performance of task 2, while Sequence 2 decreased MLT in the iterations of task 1 and task 4 (see Table 13). The decrease of MLT was restricted to specific tasks and it did not affect the MLT across tasks or across writings for any of the two WCF sequences (Table 12).
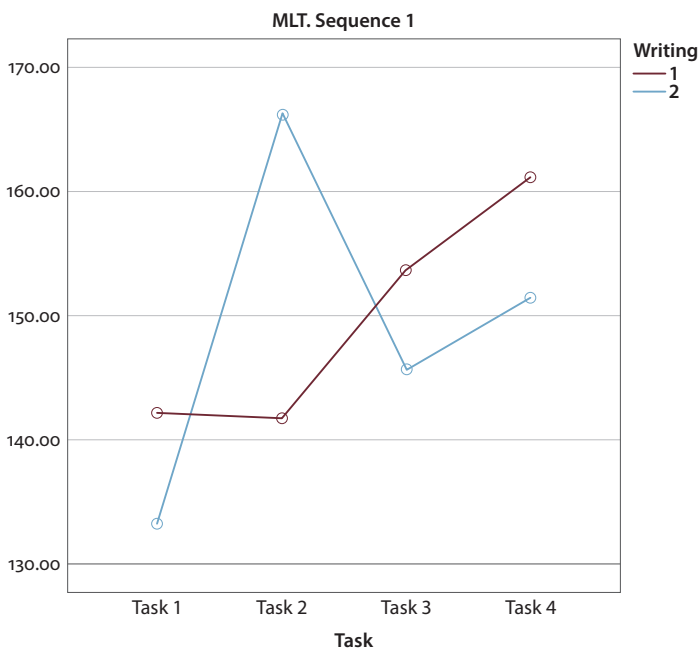
Figure 15. MLT across tasks and writings. Sequence 1

Regarding nominal complexity, both sequences increased the Mean Length of Clauses (MLC) across Tasks (Tables 12 and 13; Figures 17 and 18). In addition, the number of complex nominals per clause (CN/C) significantly increased across Tasks for both WCF sequences (Tables 12 and 13; Figures 19 and 20).
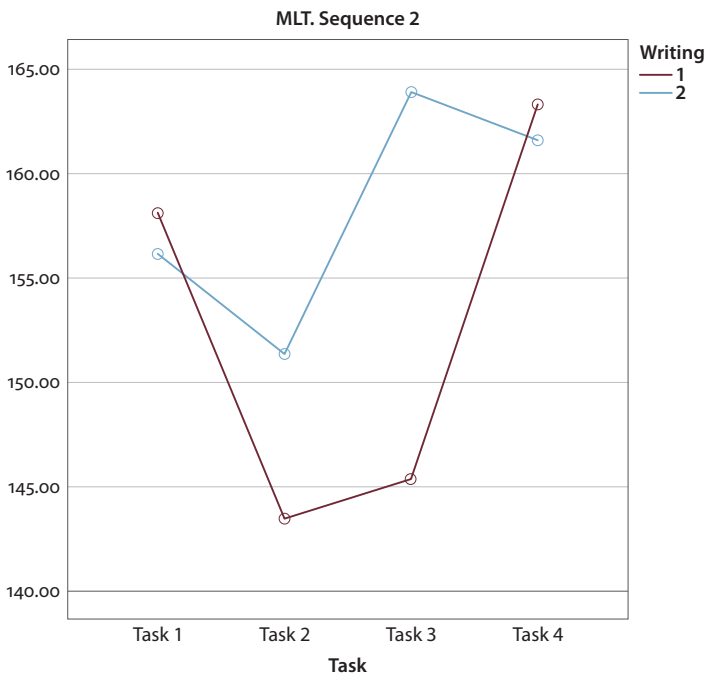
Figure 16. MLT across tasks and writings. Sequence 2
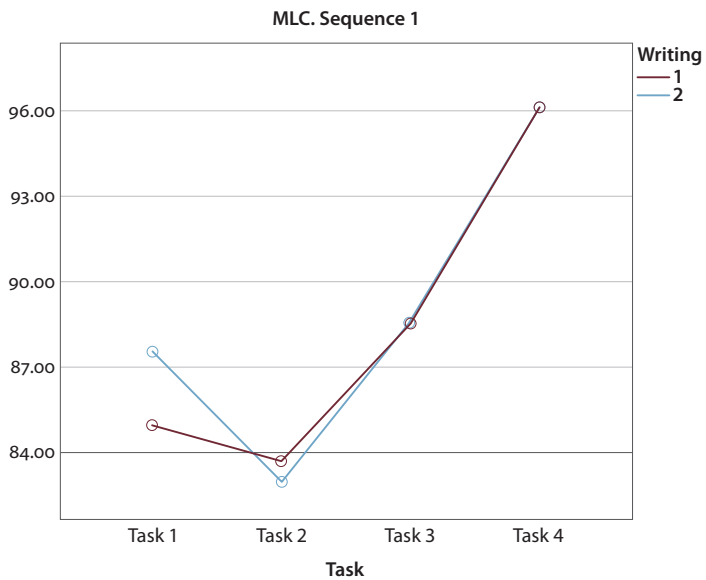
Figure 17. MLC across tasks and writings. Sequence 1
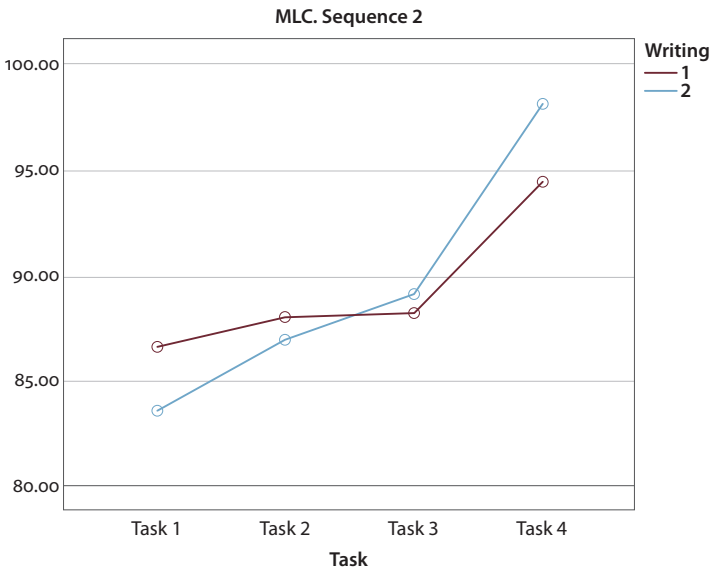
Figure 18. MLC across tasks and writings. Sequence 2
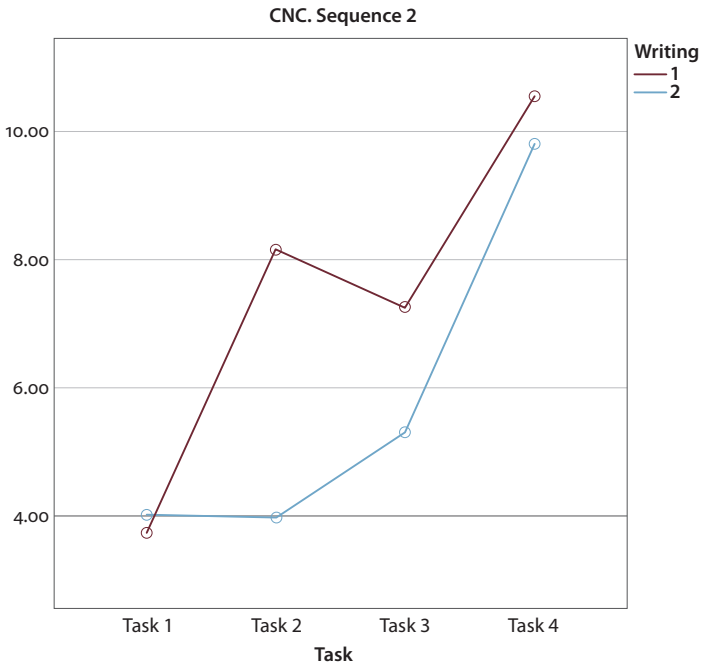
Figure 19. CN/C across tasks and writings. Sequence 1
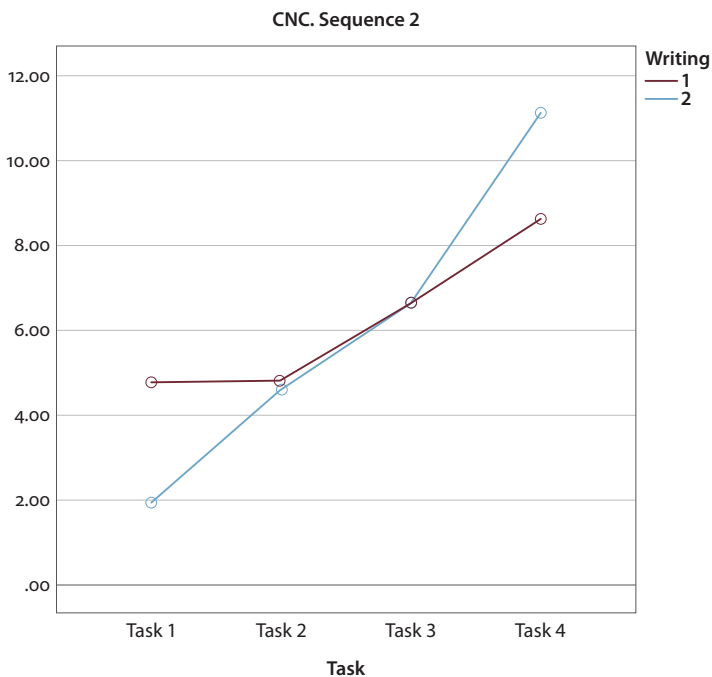
CNC. Sequence 2

Figure 20. CN/C across tasks and writings. Sequence 2

Discussion

The ultimate aim of our study was to shed further light on the learning benefits of task repetition in writing. At a global level, our findings support the beneficial effects of TR in the short- and long-term, although the dynamics of change observed shows interesting patters worth discussing from three perspectives: The language dimensions in focus (the changes observed were not isomorphic in the 3 areas of performance targeted), the temporal dimension adopted, that is, whether we are concerned with changes across iterations of the same task (where the norm was an improvement) or across time and tasks (where much more intricate patterns of development – or lack of – and a certain task-dependency of TR effects were observed), and, finally, the nature of the intervention as part of the TR cycle within and across tasks (where a conspicuous WCF-related influence on TR effects was observed). In what follows we interpret our findings from these 3 angles (we shall refer to the first 2 dimensions – areas of performance and temporal dimension – together).

Effects on performance across tasks and time

Our findings point in various directions. First, all dimensions of performance (although not all subcomponents within each CAF dimension) improved in immediate iterations of the same task and across tasks and time, the changes being more global in the area of fluency (including both total number of words and total number of sentences), and more specific in accuracy (improvements were observed only in global accuracy measures and grammar) and complexity (where TR effects were observed just in the areas of lexical complexity, and syntactic complexity in the mean length of T-units and clauses, as well as in number of complex nominal clauses). Hence, in contrast to TR oral studies (e.g. Ahmadian, 2011; Bygate, 2001; Lynch & Maclean, 2000) and to TR writing studies (especially Nitta & Baba, 2014), our study does not show trade-off effects between the three dimensions of performance either in an immediate iteration of the same task or across tasks and time. Our findings are more in line with Tavakoli's (2014) argument that the nature of writing makes it possible for learners to focus on both form and meaning simultaneously. Yet, the effects observed could be in part the result of the kind of tasks performed (demanding reading-to-write compositions), the time on task (60 minutes), and the combined effects of TR and the provision of detailed WCF, coupled with an encouragement to process it in ample time conditions and with the help of external sources or the teacher herself. It is this combination of task- and feedback-related factors, together with the relevance of writing and rewriting in the participants' learning experience, that we would suggest brought about language learning via the repetition of writing tasks over time in our study.

Second, the observed progression towards more accurate, fluent, and, in part, complex performance was not always a linear one. Thus, whereas fluency increased progressively, accuracy and complexity did not, and instead showed fluctuation across tasks and throughout the six months of the instructional program. Third, this fluctuation across tasks and time was accompanied by an equal variation in the various subcomponents of accuracy and complexity. These documented fluctuations provide strong evidence in support of Nitta and Baba's (2014) claim that the dynamics of TR effects in writing is far from being a linear, additive process.

Taken together, these findings are in part coincidental and in part divergent from those reported in previous research. At a **global level**, they appear to support Bygate's (2018) contention that "although learners' language is likely to change across iterations, we cannot confidently anticipate whether this will occur predominately in terms of complexity, or accuracy, or fluency" (p. 8). Our data allow us to add that (a) these unpredicted performance changes can go in any direction, i.e. towards improvement or towards a loss in performance (we observed

performance loss in some dimensions of accuracy and lexical sophistication), and (b) the changes that may occur can vary across time, and across subcomponents of the three macro-areas of performance targeted in the analysis. Accordingly, TR effects in writing show a distinct specificity and complexity worth exploring in new studies.

At more **specific levels**, the findings regarding the three CAF dimensions also in part coincide and in part depart from previous research. Our data distinctively point to the benefits to TR in the area of **accuracy**, which were systematically observed across task iterations and across diverse tasks and time. This finding in principle is not surprising given that mere writing practice and engaging in TR or in immediate revisions of a previously written text after receiving WCF has been systematically found to lead to more accurate writing performance (see review in Bitchener & Storch, 2016). We cannot compare our findings to previous writing TR studies as accuracy was not targeted in them (cf. Nitta & Baba, 2014), but they do confirm the benefits to task repetition in the area of accuracy reported in the oral TR literature (Bygate, 1996; Fukuta, 2015; Gass et al, 1999; Hu, 2018; Kim & Tracy-Ventura, 2013; Kobayashi & Kobayashi, 2018; Lynch & Maclean, 2000, 2001). However, in our study, the TR effects observed could be attributed to the combined effect of task iteration with plenty of time-on-task and detailed feedback provision and processing, once again in unlimited time conditions, which we would suggest is especially relevant from both research and pedagogical perspectives.

We also observed that the effects of TR varied when we looked at different error types: No substantial changes were observed in terms of lexical errors, a lineal change was observed for spelling and punctuation errors, and yet non-linear changes characterized grammar errors These findings once again point to the difficulty in anticipating precise TR effects (positive, negative or neutral, it should be stressed), as noted by Bygate (2018).

Regarding **fluency**, we corroborate the widely reported beneficial effects of TR in this performance area, in both oral and writing TR (as reviewed in Manchón, 2014a). As for the latter, our data support the increase in fluency reported in the scant research conducted on TR in the written modality (cf. Nitta & Baba, 2014. But see Chapter 6, this volume, where no effect on writing fluency was observed, although fluency was measured differently from this study). However, in contrast to Nitta and Baba's (2014) longitudinal study, in which they observed that "the benefits of specific task repetition on writing fluency were more marked in the beginning, but these benefits decreased with subsequent repetitions" (p. 118) and that TR led to higher complexity in the long term at the expense of initial increases in fluency, our data show (a) a maintained increase in fluency throughout the whole observation period, and (b) improvements in some areas of accuracy (as detailed above) and in

some dimensions of complexity (as we shall discuss next). These divergent findings can in part be explained by differences in methodology, a point we shall come back to shortly given its relevance for future research in the domain.

Finally, regarding **complexity**, the most robust TR effects (in terms of both improvement and performance loss) were observed in the longitudinal data: Over the course of the six months of the instructional program, the texts written by the participants (a) showed performance loss in lexical sophistication; (b) were equally complex in terms of coordination and subordination; and (c) improved significantly in the areas of lexical complexity, mean length of T-units and clauses, and the number of complex nominal clauses. These findings might point to the specificity and distinctiveness of TR affordances in the domain of writing and, accordingly, to the relevance of adopting methodological procedures in writing that might differ from those that are more valid for the analysis of oral performance. These methodological considerations are already part of disciplinary discussions in task-modality studies. For instance, Byrnes and Manchón (2014) considered that "this body of research [task-modality studies] is burdened by methodological problems because of *the oftentimes insufficiently critical transfer of task-as-oral-task constructs into the context of writing*. Among these are comparability of the speaking and writing tasks used, lack of sufficient control of *time on task* in order to account for the inherent time-intensiveness of writing, or *use of identical measures for analyzing both speaking and writing task performance*" (Byrnes & Manchón, 2014, p. 6. Emphasis added).

It is worth considering the divergence between our findings and those reported by Nitta and Baba (2014) from a methodological perspective as this analysis can shed light on variables worth considering in future research agendas on TR in writing that adopt a longitudinal perspective. To recall, whereas we found that TR resulted in more fluent and partly more accurate and more complex language throughout tasks and time, Nitta and Baba found that TR resulted in higher complexity in the long term in contrast to an initial increases in fluency. Yet, there are important differences between the two studies worth mentioning. One refers to the duration of the study and number of observations: six months and 4 observations points in our case versus 30 observations points over 30 weeks in Nitta and Baba's research. Hence, we may have failed to capture more subtle changes as we only had 4 snapshots of our participants' output, whereas they claim that "it is vital to have a sufficient number of data points to capture the process of changes" (Nitta & Baba, 2014, p. 28). Yet, what constitutes "sufficient number of data points" is open to interpretation. What is perhaps more relevant concerns what is and is not possible when we enter real classrooms in contrast to laboratory-type studies. In our specific case, the participants wrote regularly for their class assignments, received explicit instruction on text types for their exam preparation (we targeted

only two of those text types in the analysis) and regularly received feedback on their writing. Additionally, at four different points during the instructional period they volunteered to write and rewrite their texts in time-compressed conditions (hence reproducing the exam conditions) and these were the data that we used for our analysis. We would suggest that this is perhaps what is possible and, more important, relevant, when we conduct our research in real classrooms and do so longitudinally, that is, over the course of an entire 6-month (48 hours) curriculum. Accordingly, we would respectfully disagree with the position adopted by Nitta and Baba (2014) when they criticize the research approach of collecting "static snapshots of a writer's (or a group of writers') performance a few times and then looking for the differences between them" as, in their view, such an approach would make it "virtually impossible to capture critical changes because the timing of such changes is unlikely to correspond to that of snapshots that researchers arbitrarily choose" (p. 3).

The second difference between the two studies worth considering refers to time-on-task and task type. Thus, the participants in Nitta and Baba (2014) wrote short narrative essays in 10 minutes, in contrast to the more complex reading-to-write tasks completed during 60 minutes in our study. Additionally, our participants did not have access to external resources, whereas those in Nitta and Baba's research were allowed to use a dictionary. Most importantly, our participants had access to detailed feedback while those in Nitta and Baba's research did not. These are very different writing conditions that might explain why those in which our writers composed their texts allowed for a gradual improvement of all dimensions of their writing, whereas such TR learning outcomes might be different or more limited when TR is simply a question of successive iterations of the same task over a long period of time.

Nature of the intervention: The appropriation of WCF and its mediation in TR effects

The final lens through which we would like to discuss our findings relates to the nature of the intervention as part of the TR cycle within and across tasks, where an unpredicted but most telling WCF-related influence on TR effects was observed. Thus, the first time our participants received indirect WCF (Task 2 for Sequence 1 and Task 1 for Sequence 2) there was performance loss in the area of grammatical accuracy and mean length of T-units, two effects that did not occur the second time that both sequences received indirect WCF. This is a WCF-related effect worth discussing from the perspective of previous work on the language learning potential of direct and indirect WCF over time (Chandler, 2003; Ferris et al., 2013, Vyatkina, 2010). The general conclusion in this research is that direct WCF is more effective than other less explicit WCF types, although at times these benefits are only evident

in immediate revisions but not over time (Vyatkina, 2010), whereas in other cases longitudinal effects have been reported for indirect WCF (Ferris et al., 2013). Our results add to this complex picture: Given the WCF provision and appropriation conditions that characterized our study, the main conclusion to be drawn is to acknowledge the complexities involved in understanding the very process of WCF appropriation, the role that learner-related and instruction-related variables may play in such appropriation, and, subsequently, the evident signals in our data of the relevance of adopting a longitudinal perspective. We hope to have provided a glimpse on such complexity with the data analyzed in this chapter.

Conclusion

We would like to close with some empirical and methodological conclusions. In essence, our study adds to previous research on TR by showing the intricate patterns of writing TR effects on diverse dimensions of language and, resulting from it, to the methodological complexities involved in findings valid ways of inspecting such complexity. We interpret our data as suggesting that this complexity ought to take account of the mediation of language-related, task-related, feedback-related, and feedback-processing factors in bringing about learning via repeating writing tasks. Additionally, our findings clearly point to the need to adopt a longitudinal perspective in order to shed a stronger light on the connection between language learning and writing when this connection is investigated through the lens of task repetition.

The study is nevertheless limited given the small number of students that took part in it, the focus on just one proficiency level, the fact that we looked into a limited number of tasks and text types, and, perhaps, the limited number of data collection waves. Despite these limitations, we would like to suggest that the research reported here represents a worthy contribution to the TR literature as well to the research on the connection between writing and language learning given its focus on TR in the domain of writing (which has received much less attention than TR of oral tasks), the inclusion of an intervention between task iterations in the form of feedback (purported to be criterial in bringing about learning via TR and yet under-researched in the domain), its longitudinal nature (almost absent in previous TR work despite its claimed empirical and pedagogical relevance), the curricular framework in which it was situated (the curricular perspective thought to be critical in advancing current knowledge on writing as a site for language learning. See chapters 4, 5, and 14, this volume), and the language learners and context investigated (students of languages with no previous language/linguistics background in an out-of-school context).

Finally, our study supports the relevance of the most encompassing future research agenda on TR put forward by Bygate (2018), for whom key items in these agendas include "the extent of the range of varieties of task repetition in classroom contexts; the degree to which fluency, accuracy, complexity and other aspects of performance might each develop on different iterations; the ways in which iterations reflect changes not only in the learners' language but also in their understanding and mastery of the task" (pp. 22–23). Our research intended to be a meaningful contribution to extend current conversations by providing novel insights, and by shedding light on methodologically-valid and pedagogically-sound ways of advancing research in the field from the perspective of L2 writing and language learning.

Acknowledgements

Funding

References

Ahmadian, M. J. (2011). The effect of "massed" task repetitions on complexity, accuracy, and fluency: Does it transfer to a new task? *Language Learning Journal*, 39(3), 269–280. https://doi.org/10.1080/09571736.2010.545239

Baba, K., & Nitta, R. (2012). Dynamic effects of repeating a timed writing task in two EFL university courses. In P. M. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing: Identification, investigation and resolution* (pp. 398–413). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-60960-741-8.ch023

Baba, K., & Nitta, R. (2014). Phase transitions in the development of writing fluency from a complex dynamic systems perspective. *Language Learning*, 64(1), 1–35. https://doi.org/10.1111/lang.12033

Bardovi-Harling, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11, 17–34. https://doi.org/10.1017/S0272263100007816

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781783095056

Bulté, G., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.32.02bul

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). London: Heinemann.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 23–48). Harlow: Longman.

Bygate, M. (2006). Areas of research that influence L2 speaking instruction. In E. Uso-Juan, & A. Martinez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 159–186). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197778.3.159

Bygate, M. (2018). Introduction. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 1–25). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.intro

Byrnes, H., & Manchón, R. M. (2014). Task-based language learning: Insights from and for L2 writing. An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 1–23). Amsterdam: John Benjamins.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12, 267–296. https://doi.org/10.1016/S1060-3743(03)00038-9

Fukuta, J. (2015). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321–340. https://doi.org/10.1177/1362168815570142

Ferris, D., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22, 307–329. https://doi.org/10.1016/j.jslw.2012.09.009

Gass, S. M., Mackey, A., Álvarez-Torres, M. J., & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49(4), 549–581. https://doi.org/10.1111/0023-8333.00102

Hu, X. (2018). Effects of task-type, task-type repetition and performance criteria on L2 oral production. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 149–175). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.06hu

Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, 41, 829–840. https://doi.org/10.1016/j.system.2013.08.005

Kobayashi, E., & Kobayashi, M. (2018). Second language learning through repeated engagement in a poster presentation task. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 229–260). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.09kob

Larsen-Freeman, D. (2018). Task repetition or task iteration? It does make a difference? In M. Bygate (Ed.), *Learning language through task repetition* (pp. 317–335). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.12lar

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4, 221–250. https://doi.org/10.1177/136216880000400303

Lynch, T., & Maclean, J. (2001). Effects of the benefits of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching, and testing* (pp. 141–162). London: Longman.

Manchón, R. M. (2014a). The distinctive nature of task repetition in writing. Implications for theory, research, and pedagogy. *ELIA*, 14, 13–42. https://doi.org/10.12795/elia.2014.i14.02

Manchón, R. M. (2014b). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights to and from writing* (pp. 27–52). Amsterdam: John Benjamins.

Nicolás-Conesa, F., Manchón, R. M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new texts: Looking into feedback for accuracy and feedback for acquisition. *Modern Language Journal*, 103, 848–873. https://doi.org/10.1111/modl.12592

Nitta, R., & Baba, K. (2014). Task repetition and L2 writing development: A longitudinal study from a dynamic systems perspective. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 107–136). Amsterdam: John Benjamins.

Nitta, R., & Baba, K. (2018). Understanding benefits of repetition from a complex dynamic systems perspective. The case of a writing task. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 285–316). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.11nit

Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578. https://doi.org/10.1093/applin/amp044

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 217–236). Amsterdam: John Benjamins.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62, 1–41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Vyatkina, N. (2010). The effectiveness of written corrective feedback in teaching beginning German. *Foreign Language Annals*, 43, 671–689. https://doi.org/10.1111/j.1944-9720.2010.01108.x

Task modality, communicative adequacy and CAF measures

The moderating role of task complexity

Olena Vasylets[1], Roger Gilabert[1] and Rosa M. Manchón[2]
[1]University of Barcelona / [2]University of Murcia

In this study we looked into the learning affordances of task modalities as mediated by task complexity. A group of intermediate learners performed an argumentative task with two levels of complexity orally and in writing and their performance was rated for adequacy, and assessed in terms of the CAF (complexity, accuracy, and fluency) measures. In both oral and written modes, communicative success was found to be linked to lexical complexity and fluency. However, adequacy was associated with accuracy only in speech and with propositional complexity (idea units) only in writing. Task complexity did not moderate the links between communicative adequacy and the CAF dimensions. Based on these findings, implications for task design and for language-learning potential across modalities are drawn.

Introduction

The ability to communicate efficiently is the hallmark of a proficient L2 user. Accordingly, the aim of instructed second language acquisition (SLA) is to prepare L2 learners for successful real-life communication, which entails, *inter alia*, efficient delivery of oral and written messages that are appropriately tailored to the linguistic and pragmatic demands of the task at hand (Kuiken, Vedder, & Gilabert, 2010). These ideas form the basis of Task-based Language Teaching (TBLT), which employs pedagogic tasks modeling real-life activities as the main units of L2 syllabus, testing, and research (Ellis, 2003; Long, 2014). Notwithstanding its notable achievements, the field of TBLT has until recently been characterized by a notorious bias towards oral tasks in its theoretical and empirical research agendas (Byrnes & Manchón, 2014). Although the development of the theorizing of the language-learning potential (LLP) of L2 writing (Manchón, 2011; Manchón & Williams, 2016) has stimulated a growing interest in writing tasks, there are still

many lacunas concerning the idiosyncrasy of the learning mechanisms which operate in the written mode.

Another gap in scholarly debates is the scarcity of investigations into the construct of communicative adequacy and its relationship with the complexity, accuracy, and fluency (CAF) of production (Révész, Ekiert, & Torgersen, 2016). Although CAF constitute acknowledged measures of quality of L2 outcomes and development, they do not necessarily represent a perfect equivalent of communicative success in task performance (Pallotti, 2009). Previous research has established some patterned links between the CAF indices and communicative adequacy (Kuiken, Vedder, & Gilabert, 2010; Révész, et al., 2016), although it is still unknown whether or not the nature and strength of such relationship varies in writing and in speech, or whether task complexity influences the nature of the potential connections between CAF measures and communicative success. To fill this research gap, the study reported in this chapter explored whether the links between the CAF dimensions and communicative adequacy are mediated by the mode in which tasks are performed and by their level of complexity.

In our view, elucidating the idiosyncrasy of the connection between communicative adequacy and CAF in speaking versus writing would contribute to a better understanding of the language learning affordances of language modalities in general, and of the written mode in particular. As for the latter, the insights generated would be relevant for the theorizing on the language learning potential (LLP) of writing by shedding light on the way in which communicative success is achieved in the written mode. The comparison of communicative adequacy across language modalities could make the specific mechanisms of the construction of communicative success and the concomitant engagement of the learning mechanisms in writing even more salient. We further looked into whether these potential effects were mediated by the complexity of the task being performed, hence adding to previous SLA-oriented L2 writing research (cf. Byrnes & Manchón, 2014) on the influence of different task features (in this case, mode versus task complexity) on L2 production and learning.

In what follows we review the relevant research in order to better frame our own study in previous scholarship. We will first review and compare findings from task complexity and task modality research. We will then define the construct of communicative adequacy, and will summarize the results of previous studies exploring this performance dimension.

Task complexity and task modality in SLA research

Much of the research into tasks has adopted a psycholinguistic perspective, hence being primarily concerned with identifying those task features that can be effectively manipulated in order to promote engagement in relevant learning processes. Task complexity, defined as the intrinsic cognitive demands exerted by the task on learners' memory and attentional resources (Robinson, 2001, 2011) is one central task feature that has attracted considerable theoretical and empirical attention. The influential psycholinguistic task-based learning model known as the Cognition Hypothesis (Robinson, 2001, 2011) predicts that increasing task complexity along certain dimensions (e.g., increasing reasoning demands or the number of elements in the task) has the potential to engage L2 learners in complex thinking, creating, in this way, favorable conditions for language development.

It is interesting to note that the predictions for the purported effects of enhanced task complexity are compatible with the theoretical tenets of the LLP of L2 writing (Manchón & Vasylets, 2019, for a recent review of this research): Just as the performance of any cognitively complex task, the performance of a complex writing task, such as an argumentative essay, is supposed to effectively engage language learning mechanisms, such as enhanced attention to language or deep linguistic processing. Also, similar to the way the performance of a complex task is theorized to be more accurate and complex as compared to the performance of a simple version of the same task (Robinson, 2011), written task performance has been found to be potentially more accurate and complex than the counterpart oral task performance (Vasylets, Gilabert, & Manchón, 2017). Importantly, the problem-solving nature (i.e., inherent complexity) of many writing tasks is one of the main factors invoked when explaining the connection between writing and language learning (see Manchón & Roca de Larios, 2007).

Despite these parallelisms, task complexity theories have not included mode (oral versus written) in their taxonomies of task complexity factors (as noted by Kormos, 2014; Manchón, 2014, Tavakoli, 2014). Yet, mode has been shown to exert more stable and predictable effects on L2 performance than task complexity itself. For example, a recent study by Vasylets, Manchón, and Gilabert (2019) explored the way in which mode affected propositional and linguistic complexity and found that while written production was characterized by a higher ratio of more informationally dense ideas and higher lexical and syntactic complexity, speakers produced ideas which contained a higher number of words. The researchers concluded that, as a consequence of the slower rate of production and visibility of written output, writing constitutes a more favorable environment for the production of linguistically and propositionally complex discourse. The authors also interpreted their results as evidence of the learning outcomes of written production from the

perspective of the manner in which the production of more complex language can lead to restructuring, that is, the process by which interlanguage becomes more elaborate and structured, more likely to result in more efficient in communication, and more native-like (McLaughlin, 1990).

Even more compelling evidence for the relevance of mode in explaining performance differences and learning outcomes comes from studies exploring mode and task complexity effects within the same experimental design: Notably, these studies have consistently found that mode exerts stronger and more robust effects on L2 production than task complexity. For example, Ellis and Yuan (2005) investigated the effects of planning, mode, and their interactions on L2 narrative performance of Chinese learners of English. Although planning opportunities affected the CAF dimensions of performance, modality played a more important role, as "the extent to which…learners produced fluent, complex and accurate language depended principally on whether the task involved speaking or writing" (p. 189). Similar conclusions were drawn by Zalbidea (2017, this volume), who explored the effects of task complexity and task modality on the complexity and accuracy of production of intermediate learners of L2 Spanish. Contrary to her initial expectation, Zalbidea found that "task modality played a more robust role than task complexity in promoting improved linguistic performance among lower-intermediate learners during task-based work" (p. 348). Similarly, Vasylets et al. (2017) also found that mode exerted greater effects on performance than task complexity as written production was more syntactically complex, more lexically varied, and it contained more ideas than oral production. Another relevant finding was that written production showed more variation between the complex and simple versions of the task while oral production appeared as relatively insensitive to task complexity manipulations. The authors interpreted these findings as support of the Cognition Hypothesis when applied to writing, as well as evidence that task complexity does not operate in isolation but rather interacts with the mode in which a task is performed. Additionally, the fact that task complexity effects were primarily manifested in the written performance were also interpreted as evidence of the language learning potential of writing, which, due to its inherent characteristics, appeared to represent a perfect arena for the manifestation of the L2 learning-beneficial effects induced by increases in cognitive task demands.

In sum, previous studies provide compelling evidence that L2 performance is affected by the mode in which a task is performed. Importantly, task modality effects do not only appear to be more prominent than those of task complexity; rather, modality effects appear to operate in interaction with the increases in cognitive task demands. Moreover, recent evidence (e.g. Vasylets et al., 2017) points to the possibility of a better capacity of the written modality to channel the favorable language learning effects which potentially derive from increases in task demands.

In other words, another facet of the LLP of writing relates to the affordances of the written mode to act as a catalyzer of the learning benefits generated by task complexity.

CAF and communicative adequacy

Previous TBLT studies have most exclusively employed CAF measures to assess task complexity and task modality effects on L2 production, which is in line with the general assumption in SLA research that CAF measures should be viewed as reliable and valid indicators of L2 performance, proficiency, and development (Housen, Kuiken, & Vedder, 2012). In addition, higher CAF scores have also been taken as correlates of higher learning potential of a production mode (Ruiz-Fúnez, 2015). However, some critical voices have questioned whether CAF measures can provide a full description of L2 performance and development. Ortega (2003), for example, claimed that the development of L2 learners' discourse and sociolinguistic repertoires should be taken as another essential sign of L2 progress. Similarly, Robinson (2001) highlighted that non-linguistic, pragmatic outcomes of the task should also be evaluated in order to obtain the real estimate of successful task performance. The exclusive reliance on CAF measures was also problematized by Pallotti (2009), who defended the importance of communicative adequacy, defined as the degree to which leaners' performance is successful in achieving the communicative goal of the task. In Pallotti's conceptualization, communicative adequacy represents an independent construct that can be related to CAF measures and be used to interpret other CAF dimensions. The relationship between CAF and communicative adequacy might not be necessarily straightforward as production high on CAF may not meet genre requirements or the goals of the task. In other words, a high score of performance in terms of CAF measures might not necessarily be a guarantee of task communicative success (Pallotti, 2009).

There are several empirical studies that have explored communicative adequacy and CAF in speaking and writing. Concerning the oral mode, De Jong et al. (2012) used Structural Equation Modeling to explore the contribution of knowledge and skill variables to speaking proficiency of adult native and non-native speakers of Dutch. The results showed that, with the exception of two articulation speed measures, all linguistic skills were significantly related to speaking proficiency, with vocabulary knowledge and intonation skills being of special importance. A recent study by Révész et al. (2016) found that the main predictor of communicative success in L2 oral production was the measure of filled pause frequency, which is a sub-dimension of breakdown fluency. Révész et al. (2016) also reported that task type did not moderate the links between communicative adequacy and linguistic

features, but proficiency did in the sense that lower incidence of false starts (repair fluency) had a positive link with adequacy only for high-proficient learners.

In the case of writing, Kuiken et al. (2010) explored the relationship between communicative adequacy and CAF measures in the written production (two argumentative writing tasks of a similar type) of low-intermediate (A2-B1) learners of L2 Dutch, Italian, and Spanish. The analyses showed that adequacy scores correlated significantly with accuracy and lexical variation (Guiraud index), but not with syntactic complexity. These findings were partially confirmed by Vasylets and Gilabert (2013), who also reported that, when judging the communicative adequacy of argumentative written tasks, raters appeared to rely on text length and vocabulary (lexical sophistication, in particular), while syntactic complexity did not play a role.

It is also relevant to consider the findings in those studies that have explored the links between CAF measures and holistic judgments of global L2 proficiency. For example, Iwashita, Brown, McNamara, and O'Hagan (2008) reported that vocabulary (word type and token) and fluency (speech rate) were the most important contributors to speaking proficiency. A key finding in Sato's (2012) study with Japanese L2 English learners was that the strongest predictor of the raters' intuitive judgments of oral proficiency was the quality of content/elaboration development, followed by oral fluency. As for written production, Bulté and Housen (2014) reported that frequent use of simple sentences was perceived as an indicator of lower quality of writing, while use of longer units and use of many (and many different) words were interpreted as signs of higher quality in writing. Finally, Yang (2014) found that writing fluency (total number of words) was related to communicative success in tasks of different genres.

In sum, empirical evidence points to a link between communicative adequacy and CAF dimensions of L2 performance. However, this link is intricate and can depend on other factors, such as the level of L2 proficiency. There is also persuasive evidence of the componential nature of communicative adequacy, as various CAF dimensions have been found to contribute to functionally optimal oral or written task performance. In speech, different dimensions of fluency have been consistently linked to communicative success, although other factors, such as accuracy, vocabulary or pronunciation, have also been found to have a positive role. In writing, vocabulary and writing fluency (number of words) have emerged as important determinants of the efficiency of message transmission as well as of writing quality.

Many important issues are, however, still unresolved. Thus, in oral production the nature of the relationship between communicative adequacy and CAF measures does not seem to be moderated by task complexity. Yet, this issue has never been explored in written production and open questions also exist regarding whether or not the relationship between communicative adequacy and CAF dimensions is the same across modalities. The accumulated findings point to vocabulary as

a common feature that determines communicative success in both speaking and writing. However, the importance of the contribution of other CAF dimensions to communicative success in writing versus speech is less clear. Moreover, no previous studies have explored the manner in which propositional complexity (i.e., quantity and quality of the ideational content) may relate to communicative adequacy. It must also be mentioned that variations in the participants' characteristics and differences in tasks and instruments make it difficult to compare the nature of CAF-communicative adequacy links in the two modes across earlier studies, which motivates the present study.

The present study: Research questions

In light of the gaps in previous research discussed in the preceding sections, and in an attempt to shed light on the connection between writing and language learning from the perspective of the learning affordances of speaking and writing, our study explored and compared the relationship between CAF measures and communicative adequacy in oral versus written production. Communicative adequacy was defined as successful task completion in accordance with the instructions of the task (see elaboration in the Method section).

To facilitate the comparison between the two modes, we employed (i) the same prompts to elicit oral and written production, (ii) the same measures to assess complexity and accuracy of production, and (iii) the same scale to assess communicative adequacy in the two modes. Another specific contribution of our research is the inclusion of propositional complexity in addition to the traditional CAF measures. We also explored the links between CAF dimensions and communicative adequacy in tasks with different levels of complexity.

Accordingly, the following research questions were posed:

1. What is the relationship between communicative adequacy and CAF of L2 oral production?
2. What is the relationship between communicative adequacy and CAF of L2 written production?
3. Is the nature of the relationship (if any) between communicative adequacy and CAF dimensions different in writing and in speech?
4. To what extent does task complexity moderate the relationship between communicative adequacy and CAF in written and in oral production?

Because of the mixed nature of previous findings, we advanced no specific hypotheses and our study can thus be defined as heuristic research (Seliger & Shohamy, 1990).

Method

The study employed a repeated-measures counterbalanced design, with task complexity (simple vs. complex) as a within-subject factor, and language mode (oral vs. written) as a between-subject factor.

Participants

The participants were 78 Spanish and Catalan bilinguals, university learners of English whose ages ranged from 18 to 40 years old, with the majority being in their 20s. They were administered an X_Lex and Y_Lex vocabulary tests (Meara & Miralpeix, 2006) which were employed as a proxy of L2 proficiency. According to these tests scores, the participants had and intermediate (B1–B2) level of L2 proficiency. The participants were divided into an oral group ($n = 39$) and a writing group ($n = 39$), both groups being similar in terms of L2 proficiency, gender, and age distribution.

The experimental task

To elicit oral and written production, the Fire Chief Task (Gilabert, 2007) was employed. This task represents a schematic picture of a building on fire with various human characters. The task requires learners to verbalize and justify the most efficient plan of evacuation of the people from the building. The task has two task-complexity conditions: In the simple condition, the rescue resources are numerous and the sequence of logical actions is straightforward, whereas in the complex condition the resources are scarce and the optimal rescue scenario is less obvious. An important feature of the design of this task is that it can be administered to be performed orally and in writing without compromising its authenticity and meaningfulness. The use of the same experimental task in the two modes was relevant because the comparison of equivalent communicative tasks is an important requirement when the aim is to attempt to isolate mode-related idiosyncrasies in performance (Tannen, 1982).

Data collection procedures

The oral data were collected individually with each participant during one single session. The participant was presented with the prompt and the instructions for the simple and complex tasks. A very short planning time (for about 1 minute) was given, after which the participant had to perform both the simple and complex versions of

the task. The tasks were counterbalanced such that half of the participants performed the simple task first followed by the complex task, and the other half of the participants performed the tasks in the reversed order. The written data were collected from the participants in one group session. The procedure was similar: The participants received the prompt (counterbalanced) and they had to start writing after 1 minute of planning. No time limitations were set for task performance. Most participants took about 2 minutes for the completion of each version of the task in speaking, whereas in writing it took longer, i.e. 10–13 minutes for each task condition.

Data analysis procedures

The oral data were transcribed in CHAT (MacWhinney, 1996) and in Word, and the hand-written texts were transcribed in Word. The data were analyzed in terms of (i) objective (automated and manually-calculated) CAF measures, and (ii) subjective holistic ratings performed by non-expert raters.

CAF measures

To assess accuracy of oral and written production, the total number of errors per 100 words was calculated (all errors/words) x 100). Pronunciation errors in speech and spelling and punctuation errors in writing were not taken into consideration. For linguistic (lexical and syntactic) complexity, various measures were employed in order to account for the multifaceted nature of this dimension (Norris & Ortega, 2009). Lexical complexity was assessed using measures of lexical diversity, lexical sophistication, and lexical richness or productivity. For lexical diversity, D-value was calculated (Malvern & Richards, 2002). To measure lexical sophistication, the Lexical Frequency Profile (LFP) was calculated using the formula:

$$\text{LFP} = \frac{(\text{tokens K1} \times 1) + (\text{tokens K2} \times 2) + ((\text{tokens AWL} + \text{offlist}) \times 3))}{\sqrt{\text{tokens}}}$$

By means of Web Vocabprofile v3 (Cobb n.d.), we obtained the data about the K1 and K2 words (function and content words that are, respectively, among the most 1000 and 2000 most frequent English word families), and about the words belonging to the *Academic Word List* (Coxhead, 2000). The words which do not appear in the lists were classified as off-list. To assess lexical richness or productivity, we employed the index of Guiraud G (Guiraud, 1959). While most studies have employed *G* as an index of lexical diversity, Bulté and Housen (2014) have convincingly argued that *G* measures something more than mere diversity. The mathematical

formula of *G* overcompensates for the decrease in scores with increasing text length. As a result, the texts with higher scores for *G* do not only have fewer word repetitions, but they are also longer (Bulté, Housen, Pierrard, & Van Daele, 2008). Thus, in addition to diversity, *G* also taps into productivity, which makes it a useful complement to the D-value. Syntactic complexity was assessed in terms of general, subordinate, and phrasal complexity (Norris & Ortega, 2009). The length of analysis-of-speech (AS) units was calculated (AS-units/tokens; Foster, Tonkyn, & Wiggleworth, 2000) as a measure of general syntactic complexity. For subordination, the S-nodes per AS-unit (AS units + S-node/AS-units) was calculated. For phrasal elaboration, we obtained the mean number of modifiers per noun phrase by means of the Coh-metrix 3.0 (McNamara, Graesser, McCarthy, & Cai, 2014).

Propositional complexity was assessed in terms of (i) total number of idea units, and (ii) ratio of extended idea units (extended idea units/all idea units). And idea unit was defined as a meaningful, semantically integral chunk of discourse (Vasylets et al., 2017). Prototypical idea units are clause-like constructions. An extended idea unit represents a chunk of language consisting of the main clause and subordinate clause(s) with a strong conceptual dependency between them. Extended idea units are considered to be informationally dense, as they incorporate by default more than one semantically meaningful constituent.

Fluency was the only dimension for which we employed different measures for the oral and written productions. For writing fluency, we calculated words per minute (time/tokens) and the total number of words, which represent standard measures of product-based writing fluency. For oral fluency, we calculated indices of speed fluency: (i) Rate A, which was obtained by dividing the number of syllables produced in the whole performance by the total number of seconds of task-related speech and multiplied by 60. Rate A is calculated without pruning the production (i.e., without eliminating repetitions, false, starts and repairs); and (ii) Rate B, which was obtained by dividing the number of syllables produced in the whole performance by the total number of seconds of task-related speech and multiplied by 60. Rate B entails pruning the text (i.e., false starts, repetitions and repairs are eliminated from the performance). For the manually calculated measures (number of AS units, S-nodes, idea units and errors), 35% of data were recoded by the same rater. The Cohen's kappa coefficients ranged from 0.801 to 0.958 indicating a high degree of intra-rater reliability.

Communicative adequacy raters and scale

Based on previous work (Kuiken et al., 2010; Pallotti, 2009; Révész et al., 2016), we operationally defined communicative adequacy as successful task completion understood as relevance and efficiency of the informational content in accordance

with the instructions and genre requirements of the linguistic task at hand. The communicative adequacy of the performance was assessed by 8 raters (doctors or doctoral students in Linguistics) with no previous experience in rating production based on the holistic scale used in the study. Each rater was contacted individually and provided with the operational definition of the construct of communicative adequacy, the detailed instructions about the rating procedure, and the holistic scale with all descriptors. The raters completed their ratings in their own time. Each rater assessed spoken and written figure with the data from the performance of the simple and complex task. Each rater assessed samples of oral and written performance from both the simple and complex conditions. Each instance of spoken or written output was assessed by two raters.

The holistic rating scale used in this study was informed by previous research (De Jong et al., 2012; Kuiken et al., 2010; Kuiken & Vedder, 2014, 2017). It represented a scale in accordance with the CEFR (Council of Europe, 2001) descriptors. Following Kuiken and Vedder (2017), we attempted to elaborate a scale that would also reflect Grice's (1975) conversational maxims of quantity, relevance, manner and quality of the message transmission. Thus, our coding scheme represented a 0–6 scale (see Appendix). In this scale, zero would be assigned to a production which would not communicate any relevant information at all. The descriptors from 1–6 contained five sub-dimensions which were intended to reflect Grice's conversational maxims. The first sub-dimension in the scale descriptors was intended to reflect the maxims of quantity and relation, and it asked the raters to focus on the information units (i.e., ideas) by assessing the relevance of content, adequacy of the number of provided ideas, as well as their relatedness. The maxim of quality was gauged in the second sub-dimension which asked to assess the way task instructions and genre requirements had been fulfilled. The third sub-dimension, which focused on comprehensibility and also effort required to understand the production, was intended to connect with the maxim of manner. Finally, the scale also assessed the forth subdimension of coherence and cohesion of production, as well as the perceived general success of message transmission, which was the fifth subdimension. In sum, we strived to elaborate a theoretically and empirically motivated scale which would: (i) reflect the relevant components of communicative adequacy; (ii) provide descriptors which would be independent from the linguistic CAF measures; and (iii) serve for both oral and written modalities.

Statistical analyses

We used both descriptive statistics (means and standard deviations) and Pearson correlations, the latter to investigate the potential relationship between communicative adequacy (as assessed by raters on a six-point Likert scale) and general CAF

measures. The analysis of data normality, linearity, and homoscedasticity were performed, with no notable violations noted. Correlations were calculated separately for the simple oral, simple complex, simple written, and complex written tasks.

Results

Our first research question asked about the relationship between communicative adequacy and CAF dimensions of L2 oral production, on the one hand, and of written production, on the other. Table 1 presents the descriptive statistics for the CAF measures. At first sight, the figures look comparable over the modes and task complexity levels.

Table 1. Means and standard deviations of the CAF measures of oral and written production

	Means (Standard Deviations)												
	Errors/100 words	D-value	LFP	Guiraud	ASU length	S-nodes/AS-unit	Modifier/NP	Total idea units	Ratio of extended ideas	Words/Min.	Total words	Rate A	Rate B
Speech													
Simple condition	7,7	29,7	14,1	4,9	12,2	1,58	,76	14,4	,19	112	175	134	111
	(4,8)	(10)	(2,7)	(,84)	(2,7)	(,31)	(,11)	(4,5)	(,11)	(33)	(71)	(39)	(37)
Complex condition	7,7	31	15,5	5 (,8)	13,9	1,6	,8 (,21)	17,6	,19	110	195	129	109
	(5,2)	(9,2)	(2,3)		(3,9)	(,32)		(4,8)	(,13)	(32)	(66)	(37)	(37)
Writing													
Simple condition	7,9	44,8	14	6,3	13,4	1,91	,78	12,8	1,1	15,4	148		
	(2,76)	(13)	(1,83)	(,73)	(2,8)	(,44)	(,18)	(3,79)	(,26)	(6,76)	(61)		
Complex condition	6,5	43,9	15	6,4	14,9	2 (,43)	,82	14,7	1,2	13,1	165		
	(2,12)	(11,6)	(2,15)	(,68)	(3,3)		(,17)	(4,8)	(,26)	(4,8)	(47)		

Correlation coefficients indicated that the strength of the relationships between communicative adequacy and CAF measures in oral production ranged from medium (.412, Guiraud in the simple oral task) to large .669 (Rate B in the complex oral task). The pattern of correlations was similar in both the simple and complex conditions and can be summarized as follows: In both the simple and complex oral tasks, communicative adequacy (i) correlated negatively with the number of errors;

(ii) correlated positively with the lexical measures of D-value and Guiraud; and (iii) correlated positively with the fluency measures, including words per minute, rate A, rate B. The strength of the correlations was medium for lexical measures and large for accuracy and fluency.

Table 2. Pearson correlations between communicative adequacy as assessed by raters on a six point Likert scale and complexity, accuracy and fluency of production as assessed by general measures

Correlations	Errors/100 words	D-value	LFP	Guiraud	ASU length	S-nodes/AS-unit	Modifiers/Noun phrase	Total idea units	Ratio of extended ideas	Words/Min.	Total words	Rate A	Rate B
Bivariate													
Speech													
Simple condition	−,562**	,448**	,292	,412*	−,005	−,210	−,109	,242	−,093	,544**	,230	,531**	,589**
Complex condition	−,580**	,431*	−,226	,439*	,084	,211	,155	−,094	,165	,636**	−,207	,639**	,669**
Writing													
Simple condition	,069	,060	,558**	,427*	,065	−,196	,023	,484**	,328*	,380*	,579**		
Complex condition	−,078	,132	,502**	,296	−,111	,301	−,133	,391*	,295	,550**	,372*		

* $p \leq 0.05$
** $p \leq 0.01$

Regarding written production, and similar to the oral production, Pearson correlation coefficients ranged from medium (0.328, ratio of extended ideas, simple writing task) to large (0.579, total number of words, simple writing task). Although the pattern of correlations in the simple and complex writing tasks was not identical, there were great similarities between the two conditions (see Table 2). To summarize, communicative adequacy (i) positively correlated with lexical complexity measures, in particular with Guiraud in the simple writing task and with LFP in both the simple and complex tasks; (ii) positively correlated with the semantic measures, such as the number of ideas and the ratio of extended ideas (in the simple task); and (iii) positively correlated with the fluency measures, such as words per minute and the number of words in both the simple and complex tasks. For

lexical measures, the strength of the correlations ranged from medium (Guiraud) to large (LFP). For semantic measures, we obtained medium correlations. For writing fluency measures, the strength of the relationships with communicative adequacy ranged from medium to large, depending on the complexity condition.

Our third research question asked about the comparison of the correlations between communicative adequacy and CAF measures in speech versus writing. If we consider the complete picture of the results (Table 2), we can observe commonalities but also differences in the way communicative adequacy related to the CAF measures in the two modes. Similarities between the two modes were manifested in the presence of associations between adequacy and lexical complexity, and between adequacy and fluency, and in the absence of associations between communicative adequacy and the measures of syntactic complexity. In contrast, differences between the modes were evidenced by the fact that communicative adequacy correlated with accuracy only in speech and with the semantic dimension (idea units) only in writing.

With our final research question we intended to explore whether or not any potential links between communicative adequacy and CAF measures of oral and written production were mediated by levels of task complexity. As mentioned previously, exactly the same pattern of results was obtained in the simple and complex oral conditions. Although there were some differences between the simple and complex writing tasks (e.g., adequacy correlated significantly with Guiraud and with the ratio of complex ideas only in the simple writing condition), for the majority of measures the results were the same in the simple and complex writing conditions.

Discussion

With the ultimate aim of shedding light on the connection between writing and language learning from the perspective of the learning affordances of writing versus speaking, the study reported in this chapter explored the potential links between CAF measures and communicative adequacy in L2 oral versus L2 written production, as well as the potential mediation of task complexity in any observed effects. Accordingly, we adopted a subjective-objective approach in which holistic ratings of communicative adequacy were related to the objective measurements of accuracy, fluency, and complexity of performance. Part of the novelty of the study resides in the use of the same scale to assess communicative adequacy in the two modes. To ensure further comparability between speech and writing, the same measures (except for fluency) were used to assess oral and written performance.

Overall, the nature of the connections observed between adequacy and CAF measures was similar in speech and in writing, although intriguing differences

between the two modes were also observed. On the one hand, similarity between the two modes was reflected in the fact that lexical complexity and fluency were connected to communicative success in both oral and writing tasks, while there was no such a link for syntactic complexity. On the other hand, differences between the modes were reflected in the finding that a connection between communicative adequacy and accuracy was found only in speech. At the same time, adequacy was positively related to the quantity and quality of the ideational content only in writing.

In the discussion of these results that follows, we will start with the CAF dimensions for which the results were similar in the two modes, which will be followed by the discussion on the observed mode effects. Finally, we will discuss the role of task complexity and will draw implications for the language-learning potential of the two modes, with a specific focus on the written modality.

Similarities and differences across modalities

As mentioned above, lexical complexity was associated with communicative success in the output produced in both modes. These results are in line with those of Révész et al. (2016) regarding oral performance and Kuiken et al. (2010) or Vasylets and Gilabert (2013) in the case of writing. It is worth noting, however, that the nature of the connection between lexical complexity and communicative adequacy was somewhat different in the two modes in our data. Thus, whereas communicative adequacy was related to lexical productivity (Guiraud index) in both modes, communicative success was also connected to lexical diversity in speaking and to lexical sophistication in writing. The explanation for this variation may lie in the nature of oral and written production. Previous studies have shown that written task performance is characterized by higher lexical diversity as compared to oral production (Vasylets et al., 2017). Lower lexical diversity in speech is often attributed to the online pressures of oral production, which limits the process of lexical search and, thus, can be the cause of word repetition. Following this line of thinking, the ability to supply diverse lexis in spoken production in our data was specifically valued by our raters, who consequently associated this feature with successful performance. The absence of an association of such successful performance with lexical sophistication could be attributed to the low incidence of rare words in oral productions in our data. If not numerous, the sophisticated lexical items can go unnoticed in the fast-flowing speech and, as a consequence, lexical rarity might not capture raters' attention when assessing oral production. In contrast, the slower pace and the visibility of writing is likely to create conditions under which even few sophisticated words can be salient, and thus contribute to more positive evaluation of performance success.

Next to lexical complexity, fluency of production was also associated with communicative success in both modes. These results align with numerous speaking (e.g. Iwashita et al., 2008, Sato, 2012) and writing studies (e.g. Yang, 2014). An explanation to this finding can lie in the psycholinguistic mechanisms underlying L2 fluency. Thus, smoothness and speed of delivery of oral language have been traditionally equated with efficient and automatized access to the linguistic resources and with higher levels of development of procedural skills (Schmidt, 1992). Similarly, in writing, the number and rate of production units have been considered to be indicators of efficiency in linguistic knowledge retrieval (Wolfe-Quintero et al., 1998). Consequently, higher fluency has consistently be taken as an important determinant of successful L2 oral as well as written production.

Another common pattern across the two modes is the absence of the observable links between syntactic complexity and communicative adequacy. This result is in line with the findings by Kuiken et al. (2010) in writing, but contradicts those in Révész et al. (2016), who identified subordination as a significant predictor of adequacy in speech. These discrepancies can be attributed to the characteristics of the participants in the two studies, who were from four different proficiency levels in Révész et al. (2016) as compared to intermediate levels in the present study. There is evidence that the connection between CAF dimensions of performance and communicative success may vary as a function of L2 learners' proficiency level (Iwashita et al., 2008). Higgs and Clifford (1982), for example, theorized that grammar would play the most dominant role in overall speaking proficiency at the lowest level. Applying this idea to both modes, we could suggest that the relative weight of syntactic complexity played a less salient role in oral and written productions of our subjects as they possessed an intermediate (B1–B2) level of L2 proficiency.

To sum up, the associations between communicative success and CAF measures were of similar nature for syntactic and lexical complexity and fluency in both modes. There were, however, two core CAF dimensions for which the links with communicative success differed depending on the mode of production. Thus, a positive association between accuracy and communicative adequacy was found only in speech (see also Révész et al., 2016), while the semantic dimension of performance (i.e., quantity and quality of idea units) was related to communicative success only in writing. To explain these finding, we could again draw on some inherent features differentiating oral and written discourse. As mentioned above, speech is faster and it is also characterized by pronunciation and other prosodic features, such as voice quality, rhythm, or intonation. The participants in or study were at the low- and high-intermediate level of L2 proficiency, which can be characterized by low and irregular fluency, as well as showing nonnative rhythm and intonation. It is plausible that the flaws in these features made linguistic errors even more salient, with a consequent connection of higher accuracy to successful performance. In

contrast, lack of influence of prosodic features in writing could have created the conditions under which the influence of linguistic errors becomes less salient and sematic content acquires more importance for raters. It is difficult to compare our findings with those of previous research because of the scarcity of investigations exploring the role of content dimension in L2 communicative success. A notable exception is a study by Sato (2012), who found that content elaboration, defined as the degree to which learners conveyed relevant and well-developed ideas, made a major contribution to adequacy of L2 speech. Findings in our study only partially confirm Sato's (2012) results, as we found this effect for written but not for oral production. One potential explanation of this discrepancy could be the differences in the tasks employed in both studies. While in our study we used an argumentative task, the participants in Sato's (2012) study performed both an argumentative and a descriptive essay, which could have elicited different patterns of performance. Differences in the participants' L2 proficiency profiles could be another explanation, intermediate (B1–B2) in our case, while in Sato's research L2 proficiency ranged from intermediate to advanced levels.

Finally, our results showed that task complexity did not influence the associations between communicative adequacy and CAF measures in the oral or written mode to any substantial degree. These results are in line with those in Révész et al.'s (2016), who reported that task type did not moderate the relationship between adequacy and CAF measures. In our study, in both the simple and complex oral tasks, communicative success was linked to lexical productivity and diversity, accuracy, and speed fluency. No substantial differences were attested between the simple and complex writing tasks. Thus, in both writing conditions, we found a connection between communicative success and lexical sophistication, idea units, and fluency. This allows us to conclude that task complexity did not moderate, in any substantial way, the links between communicative success and CAF measures neither in oral nor in written production.

Implications of the findings for the connection between L2 writing and L2 learning

Our results can be interpreted from the perspective of the idiosyncratic nature of the language learning opportunities in the two modes of production. Thus, the finding that communicative adequacy was associated with different dimensions of lexical complexity in speech as compared to writing could point to the possibility that oral and written modes may differ in their affordances for L2 lexical development. The fast and evanescent nature of speech may put greater limits on lexical search and monitoring processes, resulting in impoverished lexical diversity and/ or lexical sophistication in performance (Ellis & Yuan, 2005; Vasylets et al., 2017).

In contrast, the slower pace of writing and the visibility of written output can create more favorable conditions for lexical retrieval, hence increasing the likelihood of writers making full use of their L2 mental lexicon. Future research, including both classroom-based and more controlled studies, should further explore the possibility of a differential nature of L2 lexical development in the two modes.

Another result that we must highlight is the salient role that the semantic dimension appeared to play in achieving communicative success in written production. We must bear in mind that a focus on meaning (i.e., semantic content) is a primary concern in communicative language teaching. According to our results, writing has a potential to offer favorable conditions under which a focus on meaning might be reinforced, which is reflected in our data in the link between the quality/quantity of semantic content and communicative success. This shows that the learning affordances of the written mode could align with the learning aims of communicative language teaching, which attributes special prominence to holistic and meaningful language use. This entails, inter alia, efficient conveyance of semantic content.

Another notable result worth exploring from the perspective of the learning potential of writing was that adequacy was connected to speed fluency in both modes, which is considered an index of effectiveness of language access and retrieval processes (Wolfe-Quintero et al., 1998). We must bear in mind, however, that the speed of performance is inherently different in speech and writing. Production of oral language is faster and it takes place under online pressure, and these characteristics of speech create a condition under which the speed of processing becomes fundamental. Additionally, as shown in previous studies, speed of language production constitutes a core factor of automaticity, which is an important learning process (DeKeyser, 2007). Hence, another key concern for future research should be to elucidate whether or not writing conditions, which are inherently slower, contribute to the development of automaticity.

Importantly, we also found that the nature of the associations between communicative success and CAF measures depended more on the mode of production than on task complexity. This finding adds to the growing empirical evidence showing that mode of performance exerts more robust and potentially more predictable effects on L2 learners' performance than task complexity. In turn, this confirms the potential of mode as a powerful task design variable to be taken into consideration when designing effective language tasks. The available evidence attests to the potential of writing tasks to elicit more complex production as compared to oral tasks (Ellis & Yuan, 2005; Vasylets et al., 2017, 2019). Future research, however, should overcome the prevalent focus on the outcomes of task performance and put more emphasis on the investigation of learning processes engaged in during oral versus written task performance (Révész, 2013. See also Leow & Manchón, and Zalbidea,

this volume). More focus on processes would, undoubtedly, contribute to shedding a stronger light on the singularity of the way in which oral and written modes engage language learning mechanisms.

Conclusion

In this study we employed the same tasks and the same scale to elicit and assess language production in speech and in writing. This allowed us to uncover new intricacies of the links between communicative success and CAF measures in the two modes. We confirmed some previously reported findings, such as a connection between communicative adequacy and vocabulary, and also between adequacy and fluency in both modes. However, we also found that communicative adequacy was connected to accuracy only in speech, while semantic content was visibly linked to communicative success only in writing. These findings did not only allow us to reach conclusions about the singularity of the way communicative success is construed in the two modes, but they do provide a basis for inferences about the language learning potential of the two modes. Finally, in contrast to mode, task complexity did not exert any substantial influence on the associations between adequacy and CAF of production in speech or in writing. We interpret this finding as evidence for the robustness of the effects of mode on L2 performance and (potentially) processing as compared to those of task complexity.

To conclude, a number of limitations must also be acknowledged. In the first place, we used only one type of task and we explored the relationship between communicative adequacy and general measures of production. In future studies it would be interesting to investigate other types of tasks (e.g., a narrative task) and to complement general measures with specific measures, which could be particularly relevant for the successful completion of given tasks. Future studies with a greater number of participants from different learning contexts, of different ages, and with varying levels of L2 proficiency would be desirable. Our study could also have benefited from interviewing raters about their perceptions and decisions when assessing adequacy of production.

Despite these limitations, we consider that our study has made a contribution to previous work on the connection between L2 writing and language learning by providing new insights on the nature of communicative success across modalities and the singularity of learning opportunities in speech and writing.

Acknowledgements

References

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Bulté, B., Housen, A., Pierrard, M., & Van Daele, S. (2008). Investigating lexical proficiency development over time–the case of Dutch-speaking learners of French in Brussels. *Journal of French Language Studies*, 18(3), 277–298. https://doi.org/10.1017/S0959269508003451

Byrnes, H., & Manchón, R. M. (Eds.). (2014). *Task-based language learning. Insights from and for L2 writing*. Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.7

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. https://doi.org/10.2307/3587951

Council of Europe. (2001). *Common European Framework of references for languages: Learning, teaching, assessment*. Cambridge: CUP.

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5–34. https://doi.org/10.1017/S0272263111000489

DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511667275

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: OUP.

Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–193). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.11ell

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, 21(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *IRAL-International Review of Applied Linguistics in Language Teaching*, 45(3), 215–240. https://doi.org/10.1515/iral.2007.010

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Speech acts* (pp. 41–58). New York, NY: Academic Press. https://doi.org/10.1163/9789004368811_003

Guiraud, P. (1959). *Problèmes et methods de la statistique linguistique*. Dordrecht: Reidel.

Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy, and fluency in SLA*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.32

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29(1), 24–49. https://doi.org/10.1093/applin/amm017

Kormos, J. (2014). Differences across modalities of performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 193–217). Amsterdam: John Benjamins.

Kuiken, F., & Vedder, I. (2014). Rating written performance: What do raters do and why? *Language Testing*, 31(3), 329–348. https://doi.org/10.1177/0265532214526174

Kuiken, F., & Vedder, I. (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing*, 34(3), 321–336. https://doi.org/10.1177/0265532216663991

Kuiken, F., Vedder, I., & Gilabert, R. (2010). Communicative adequacy and linguistic complexity in L2 writing. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 81–100). Eurosla Monographs Series 1.

Long, M. (2014). *Second language acquisition and task-based language teaching*. New York, NY: John Wiley & Sons.

MacWhinney, B. (1996). The CHILDES system. *American Journal of Speech-Language Pathology*, 5(1), 5–14. https://doi.org/10.1044/1058-0360.0501.05

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85–104. https://doi.org/10.1191/0265532202lt2210a

Manchón, R. M. (Ed.). (2011). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31

Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 27–53). Amsterdam: John Benjamins.

Manchón, R. M., & Roca de Larios, J. (2007). Writing-to-learn in instructed language learning contexts. In E. A. Soler & M. P. S. Jordá (Eds.), *Intercultural language use and language learning* (pp. 101–121). Berlin: Springer. https://doi.org/10.1007/978-1-4020-5639-0_6

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge: CUP. https://doi.org/10.1017/9781108333603.015

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Meara, P. M., & Miralpeix, I. (2006). *Y_Lex: The Swansea advanced vocabulary levels test* (Version 2.05)[Computer software]. Swansea: Lognostics.

McLaughlin, B. (1990). Restructuring. *Applied linguistics*, 11(2), 113–128. https://doi.org/10.1093/applin/11.2.113

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511894664

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. https://doi.org/10.1093/applin/amp044

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518. https://doi.org/10.1093/applin/24.4.492

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590–601. https://doi.org/10.1093/applin/amp045

Révész, A. (2013). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics*, 35(1), 87–92. https://doi.org/10.1093/applin/amt039

Révész, A., Ekiert, M., & Torgersen, E. N. (2016). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics*, 37(6), 828–848.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27–57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 3–39). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.2.05ch1

Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing*, 28, 1–19. https://doi.org/10.1016/j.jslw.2015.02.001

Sato, T. (2012). The contribution of test-takers' speech content to scores on an English oral proficiency test. *Language Testing*, 29(2), 223–241. https://doi.org/10.1177/0265532211421162

Seliger, H. W., & Shohamy, E. (1990). *Second language research methods*. Oxford: OUP.

Tannen, D. (1982). The oral/literate continuum in discourse. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 1–16). Norwood, NJ: Ablex.

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 217–236). Amsterdam: John Benjamins.

Vasylets, O., & Gilabert, R. (2013). Communicative adequacy: The neglected component in L2 writing assessment. Paper presented at the conference of Urban Multilingualism & Education, University of Ghent, Belgium, March 7–8.

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394–430. https://doi.org/10.1111/lang.12228

Vasylets, O., Gilabert, R., & Manchón, R. M. (2019). Contribution of oral and written modes to lexical, syntactic and propositional complexity in L2 performance in instructed contexts. *Instructed Second Language Acquisition*, 3(2), 206–227.

Wolfe-Quintero, K., Inagaki, S., Kim, H. Y., Kim, H. Y., & Inagaki, S. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (p. 187). Honolulu, HI: University of Hawai'i Press.

Yang, W. (2014). Mapping the relationship among the cognitive complexity of independent writing tasks, L2 writing quality, and complexity, accuracy and fluency of L2 writing (Unpublished doctoral dissertation). Georgia State University. Retrieved from <https://www.semanticscholar.org/paper/Mapping-the-Relationships-among-the-Cognitive-of-L2-Yang/8f180e9c319f5a700a646bed4aea1f3fd969e7ad> (26 May, 2020).

Zalbidea, J. (2017). 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *Modern Language Journal*, 101(2), 335–352. https://doi.org/10.1111/modl.12389

Appendix

0	1	2	3	4	5	6
The participant does not communicate any relevant information	The *ideas* are *inadequate* and *unrelated* to each other. The number of ideas is evidently insufficient.	The *ideas* are *scarcely adequate* and *scarcely related*. The number of ideas is insufficient.	The *ideas* are *inadequate* and *unrelated* to each other. The number of ideas is insufficient.	The *ideas* are *somewhat inadequate* and *a bit unrelated* to each other. The number of ideas is somewhat sufficient.	The number and the content of *ideas* are *very adequate.* The ideas are *related* to each other.	The number and the content of *ideas* are *extremely adequate.* The ideas are *very much related* to each other.
	None of the *requirements* of an argumentative task are met.	*Some (less than half)* of the *requirements* of an argumentative task are met.	*Approximately half of* the *requirements* of an argumentative task are met.	*Most (more than half)* of the *requirements* of an argumentative task are met.	*Almost all* of the *requirements* of an argumentative task are met.	*All* of the *requirements* of an argumentative task are met.
	The text is *not at all comprehensible*. It takes *a lot of effort* to understand the production.	The production is *scarcely comprehensible*. It takes *effort* to understand the production.	The production is *somewhat comprehensible*. It takes *some effort* to understand the production.	The production is *comprehensible*. It takes *no effort* to understand the production.	The production is *easily comprehensible*.	The production is *very easily comprehensible*.
	The production is *not at all coherent/ cohesive*: there are numerous coherences breaks, very few connectives are used, the production is *very confusing*.	The production is *scarcely coherent/ cohesive*: there are coherences breaks, few connectives are used, the production is *confusing*.	The production is *somewhat coherent/ cohesive*: there are some coherences breaks, more connectives could be used, the production is *somewhat confusing*.	The production is *coherent/cohesive*: coherences breaks are rare, use of connectives is rather appropriate, the production is *not confusing*.	The production is *very coherent/ cohesive*: there are no coherence breaks, use of connectives is appropriate, the production is *comprehensive*.	The production is *extremely coherent/ cohesive*: there are no coherence breaks, use of connectives is very skilful, the production is *highly comprehensive*.
	This is an *unsuccessful contribution*.	This is a *weak* contribution.	This is a *moderately successful* contribution.	This is a *successful* contribution.	This is a *very successful* contribution.	This is a *highly successful* contribution.

A mixed-methods approach to exploring the L2 learning potential of writing versus speaking

Janire Zalbidea
Temple University

The study reported in this chapter investigated the extent to which task modality can impact L2 learners' (a) noticing of the target form, (b) target form incorporation, and (c) perceptions of task-induced demands in a grammar-focused task. Using a mixed-methods approach, three sources of information were examined: (a) stimulated recall protocols, (b) learners' task performance, and (c) subjective post-task questionnaires. Although both task modality conditions led to substantial noticing and form incorporation, along with similar ratings of task demands, participants in the writing condition were more accurate in incorporating the target form into their own output. Findings are discussed in terms of their implications for the L2 learning potential of writing versus speaking and the contributions of mixed-methods approaches to the study of task modality in SLA.

Introduction

From a task-based perspective, output/input modality has been posited to influence learners' opportunities for L2 development (e.g., Gilabert, Manchón, & Vasylets, 2016; Williams, 2012). Specifically, the written modality has been hypothesized to promote the engagement of L2 learning processes, such as noticing and deeper processing of target input, to a greater extent than the oral modality. This prediction rests on the modality-inherent features of written compared to oral communication (e.g., slower rate of production, visual nature of input and output), which are expected to provide psycholinguistic advantages for L2 development (e.g., Gilabert et al., 2016; Manchón, 2014; Williams, 2012).

Despite increased interest in researching task modality, evidence for the differential L2 learning potential of writing versus speaking is still limited (see Gilabert et al., 2016), particularly from studies that consider quantitative and qualitative information from multiple data sources. In order to contribute to the goal of the

second part of this volume, the study reported in this chapter sought to investigate the extent to which task modality impacts learners' opportunities for L2 grammar learning as indicated by potential differences in learners' noticing, form incorporation, and perceptions of task demands in written versus oral task conditions.

In what follows, I summarize relevant theoretical background pertaining to task modality and review prior empirical findings on the role of output/input modality for promoting L2 development, particularly as they relate to the three data sources considered here. Next, I report on the methods and results of this mixed-methods study investigating the effects of task modality.

Task modality and L2 learning potential

The field of second language acquisition (SLA), particularly within the Task-based Language Teaching (TBLT) framework, has long been interested in researching the potential of different task designs for promoting L2 development. One task design component that has received increased attention in recent years is that of L2 task modality, that is, whether learners complete the task via written or oral output and input. From a psycholinguistic perspective, the act of language production and its resulting output differ widely in each modality. On the basis of these differences, several authors have hypothesized that the written modality may provide certain advantages over the oral modality for promoting SLA processes (e.g., Gilabert, Manchón, & Vasylets, 2016; Manchón, 2014; Williams, 2012).

From a theoretical standpoint, writing and speaking are both characterized as involving several production processes. According to Kellogg (1996 and elsewhere), writing is a recursive activity that involves three primary processes: formulation, execution, and monitoring. During formulation, writers plan and shape their message and also translate that message into its corresponding verbal structure. Execution entails transcribing the verbal structure of the message into text via motor skills. Lastly, when writers engage in monitoring, they read the text and identify any inaccuracies that might be found in the actual or planned discourse, revising them accordingly. Speaking also comprises various major stages, according to Levelt's (1989 and elsewhere) model of language production: conceptualization, formulation, and articulation. Speakers first engage in conceptualization by determining and organizing their message. Next, speakers turn the message into verbal code by way of grammatical and phonological encoding. Finally, the verbal material is articulated in the form of overt speech. Speakers also monitor their verbal and preverbal message, engaging in self-repair as needed.

Although writing and speaking draw on analogous mechanisms for language production (Cleland & Pickering, 2006), they differ along a number of key

dimensions concerning the act of production and its final outcome (e.g., Gilabert et al., 2016; Kormos & Trebits, 2012; Kuiken & Vedder, 2012; Ortega, 2012; Vasylets, Gilabert, & Manchón, 2017; Williams, 2012). Specifically, speaking is faster than writing and bears higher temporal pressure to produce output online. In contrast, writing proceeds at slower rates, and writers have the option of backtracking and also bear greater control over language production processes. Additionally, oral language is ephemeral and phonic, whereas written language is long-lasting and visual (see Gilabert et al., 2016). In light of these important differences, and as part of the more general debate on the connection between writing and language learning, there has been a call to further investigate the "language learning potential" (Manchón, 2014, p. 44) that the written and oral modalities afford in the context of a task.

In this respect, following Leow's (2015) model of SLA, Gilabert et al. (2016) propose that the written and oral modalities may offer different opportunities to engage in L2 learning processes in task-based contexts. They note that, in the written modality, "the permanence and self-paced nature of processing input supplied to learners may liberate attentional resources that may facilitate noticing" (p. 123) and deeper processing of L2 forms. Thus, Gilabert and colleagues argue that the written modality can facilitate SLA processes to a greater extent than the oral modality in certain conditions. Their hypothesis is consistent with earlier claims in the literature about the superiority of writing for promoting learners' attention to form (e.g., Williams, 2012). In what follows, I review prior empirical findings on the role of task modality for promoting SLA processes and outcomes.

Prior empirical research on task modality and SLA

Even though both modalities are prevalent across L2 learning contexts, task modality has received little attention in empirical SLA research, at least in comparison to other task-related factors. Within TBLT research, the role of task modality has been investigated mostly in relation to learners' L2 production performance (e.g., Kormos & Trebits, 2012; Vasylets, Gilabert, & Manchón, 2017; Zalbidea, 2017). However, research examining the extent to which task modality affects learner-generated noticing, incorporation, or task perceptions from an L2 development perspective is scarce (Gilabert et al., 2016).

In general, earlier interactionist research has identified certain advantages for writing over speaking in support of learners' focus on L2 form, although these benefits are not consistently observed across studies. For instance, Niu (2009) found that collaborative text reconstruction tasks generated more discussion during the resolution of language-related episodes in the written modality compared to the

oral modality. Similarly, Adams and Ross-Feldman (2008) showed that collaborative writing tasks promoted greater discussion of L2 forms than speaking tasks, although these differences were only observed at a descriptive level. Adams (2006) also found an advantage for writing compared to speaking in terms of learner attention to L2 form during information-gap tasks, in line with the theoretical postulations above. However, a more recent study by Rouhshad, Wigglesworth, and Storch (2016) failed to find greater negotiations for form in interactive written tasks carried out in text-chat compared to oral tasks conducted in face-to-face conditions. As suggested by Rouhshad and colleagues, this difference in findings may be due to the fact that the more meaning-oriented focus of their interactive task promoted learners' focus on form to a lesser extent than some of the tasks used in prior research.

Research focusing on the role of feedback for L2 development has also resulted in divergent findings for modality. Whereas Sheen (2010) found that both written and oral metalinguistic feedback provided on narrative tasks were equally effective in promoting L2 grammar development, Sagarra and Abbhul (2013) showed that aurally enhanced recasts were more beneficial than both visually enhanced recasts as well as unenhanced written and aural recasts when learners completed a computerized fill-in-the blank grammar task. Furthermore, within computer-mediated interaction research, Baralt (2013) reported that recasts were equally effective in the oral face-to-face modality versus the written text-chat modality for L2 grammar development among intermediate learners. Nonetheless, oral recasts promoted greater learning when the cognitive complexity of the task was increased. Conversely, other research has found that providing feedback in the form of recasts (Yilmaz & Yuksel, 2011) or either recasts or explicit feedback (Yilmaz, 2012) via written text-chat was more effective than oral face-to-face feedback for L2 grammar learning among *ab initio* learners.

Overall, results from prior studies focusing on L2 interaction (e.g., language-related episodes) or language development have identified a relevant role for task modality. Although research does point to certain advantages for the written modality over the oral modality, these differences appear to be contingent on a number of key methodological aspects, including type of task and type of feedback. Importantly, when modality differences are identified, findings are generally explained on the basis of theoretical assumptions about how modality might have affected SLA processes, such as noticing, during L2 instruction (see Gilabert et al., 2016). However, in order to better understand how modality actually influences L2 learner behavior, and thus assess how modality constrains a task's "language learning potential" (Manchón, 2014, p. 44), it would be highly informative to gather data that are indicative of these processes. With this goal in mind, a relatively small number of recent studies have sought to examine how modality might affect learners'

noticing (Gurzynski-Weiss & Baralt, 2015; Lai & Zhao, 2006; Yuksel & Inan, 2014; Ziegler, 2017) and form incorporation into their own output (Gurzynski-Weiss & Baralt, 2014; Ziegler, 2017), as well as their perceptions of task demands (Baralt, 2013; Cho, 2018). These studies are reviewed in the following section.

Indicators of L2 learning potential: Noticing, incorporation, and task perception

Indicators of L2 learning potential refer here to sources of information that signal the opportunities for L2 learning afforded by a given task. Within task-based language learning parameters, three major indicators stand out as particularly relevant: (a) the extent to which learners engage in *noticing* of the target form; (b) the quantity and quality of learners' *incorporation* of the target form into their own output; and (c) learners' *perceptions* of task-generated demands. Indeed, research has shown that noticing (i.e., the conscious cognitive registration of a form) is a necessary (albeit not a sufficient) condition for L2 development to take place (Schmidt, 2001), and that engaging in noticing and deeper processing of forms results in greater and more robust learning (Leow, 2015). Similarly, as a result of noticing, learners may engage in form incorporation, whereby they integrate the target form noticed in the input into their own output, a process that is posited to be beneficial for L2 development (Robinson, 2001, 2003). One way in which learners may incorporate forms in ways shown to benefit SLA is by modifying their output immediately after receiving feedback (e.g., McDonough, 2004, 2005). Lastly, tasks in which learners engage in deeper processing and form incorporation require learners to actively deploy their attentional resources and exert substantial cognitive effort, and may thus be perceived as more demanding or difficult by learners themselves. In light of these considerations, a task that promotes greater noticing and form incorporation, and that is perceived as more demanding by learners (e.g., Robinson, 2001), can be hypothesized to provide greater L2 learning potential.

Indeed, empirical research has shown that task modality can influence learners' noticing behaviors (Lai & Zhao, 2006; Yuksel & Inan, 2014; Ziegler, 2017), and thus impact their opportunities for subsequent L2 development. In an early study, Lai and Zhao (2006) showed that text-chat conditions were more conducive to learner noticing than face-to-face interactive conditions, although this was the case for learners' noticing of their own output but not for noticing of input (feedback in the form of recasts). Extending this study, Yuksel and Inan (2014) found that learners engaged in greater noticing of negotiation for meaning in the text-chat condition compared to the face-to-face condition. More recently, however, Ziegler (2017) and Gurzynski-Weiss and Baralt (2015) identified no substantial differences in terms of noticing across different task modalities. Overall, current evidence does suggest

that the written modality may have greater potential than the oral modality to promote learner-generated noticing in certain task conditions, although these modality benefits are not always borne out across studies.

Above and beyond noticing, research has also investigated the role of task modality on learners' incorporation of input, specifically regarding the quantity and quality of their modified output immediately following the provision of feedback (Gurzynski-Weiss & Baralt, 2014; Ziegler, 2017). Gurzynski-Weiss and Baralt (2014) compared learners' performance on face-to-face and text-chat conditions and found that learners produced less modified output in the text-chat condition. However, partial modified output was indicative of greater noticing in the text-chat condition, but not in the face-to-face condition (Gurzynski-Weiss & Baralt, 2015). Conversely, Ziegler (2017) found that learners produced comparable levels of modified output across oral tasks (face-to-face and video-chat) and written tasks (text-chat), as did Lai and Zhao (2006). Thus, current findings regarding form incorporation – more specifically, as it pertains to modified output – appear to reveal no clear benefits for one modality over the other.

Lastly, a relatively small amount of research has also explored whether written and oral tasks generate different cognitive demands as measured by learners' ratings in subjective task perception questionnaires. Although, as noted earlier, Baralt (2013) found no major differences in the extent to which written and oral interactive tasks promoted L2 development, differences were identified in terms of learners' perceptions of task demands. Specifically, oral tasks that required greater reasoning led to higher ratings of task difficulty and anxiety than written tasks. A more recent study by Cho (2018) found, in line with Baralt (2013), that task modality affected learners' perceived task difficulty, competence, and balance. Writing was perceived as less difficult than speaking, and learners reported greater competence in writing than speaking. These results led Cho to claim that the written modality may afford more positive learner experiences than the oral modality.

Taken together, findings from earlier research indicate that, consistent with the theoretical predictions outlined above, tasks in the written modality can promote greater noticing and may also be perceived as less difficult by L2 learners. Findings regarding incorporation in the form of learner modified output are less clear cut, however, in revealing differences on the basis of task modality. As suggested by Ziegler (2017), inconsistent results across studies may be in part due to methodological differences in task design, as well as in the operationalizations of noticing and modified output. It is important to highlight here that, in seeking to explore the L2 learning potential of each modality, prior research has considered only one or two of these indicators (noticing, incorporation, or learner perceptions) within the same study, which may provide a partial view into the affordances of each modality. By considering multiple indicators and combining quantitative and qualitative

methods using a mixed-methods approach (e.g., Riazi, 2016), researchers can more reliably assess the potential of task modality and better understand how its effects are manifested in learners' L2 performance and cognition. This is the goal of the present study.

The study

Research question

The study addressed the following research question: *To what extent does task modality (writing vs. speaking) impact L2 learners' (a) noticing, (b) form incorporation, and (c) perceptions of task demands in an L2 grammar-focused task?*

Three data sources were considered: (a) stimulated recall protocols were employed to evaluate noticing, (b) learners' performance in a computerized focused task was examined for instances of form incorporation, and (c) subjective post-task questionnaires were administered to explore learners' perceptions of task demands.

Method

Participants

Participants were 55 beginning learners of L2 Spanish who were randomly assigned to a Writing group ($n = 28$, 15 female) or a Speaking group ($n = 27$, 20 female).[1] Forty-nine participants reported English as their L1 (4 participants were native speakers of Romanian, Greek, Arabic, and Patois in addition to English). Six participants indicated a different L1 (Chinese, Turkish, Amharic, Urdu, and Bahasa Indonesia), although they reported an early age of exposure to English ($M = 4.6$ years, $SD = 1.2$). At the time of the study, all participants were enrolled in first-semester (Writing: $n = 24$; Speaking: $n = 22$) or second-semester (Writing: $n = 4$; Speaking: $n = 5$) Spanish language courses at the university. These courses were part of a communicative language teaching-based program. All participants

---

1.   Participants were excluded from the initial sample ($N = 63$) if (a) their age of exposure to Spanish was ≤ 3 years old ($n = 1$), (b) they did not follow focused task instructions ($n = 2$), (c) they produced the target structure accurately from the beginning of the focused task ($n = 1$), or (d) their performance on pretest assessments was above established cut-off points ($n = 4$). One participant was excluded from the post-task ratings dataset because they failed to complete the questionnaire immediately after the focused task.

completed a language background questionnaire where they reported their age of exposure and years of formal education in Spanish, among other information (see Table 1). No statistical differences were found between groups with regard to the background variables: age, $U = 329.000$, $z = -.880$, $p = .379$; number of foreign languages, $U = 349.000$, $z = -.570$, $p = .569$; formal education in Spanish, $U = 363.000$, $z = -.258$, $p = .797$; age of exposure to Spanish, $U = 361.000$, $z = -.288$, $p = .773$, or self-rated Spanish proficiency, $U = 314.500$, $z = -.651$, $p = .515$.

Table 1. Participant background information

	Writing		Speaking	
	M (SD)	Mdn	M (SD)	Mdn
Age at time of study	19.11 (1.23)	19.00	18.81 (1.04)	19.00
Number of foreign languages	1.39 (.57)	1.00	1.56 (.80)	1.00
Formal education in Spanish (years)	3.45 (3.25)	2.50	3.11 (3.02)	2.50
Age of exposure to Spanish	14.82 (4.56)	15.00	14.52 (3.98)	15.00
Overall self-rated Spanish proficiency[a]	4.04 (1.87)	4.00	3.56 (1.48)	3.63

Note. Scores represent ratings on a 10-point Likert scale averaged across all four skills.

Target structure

The target structure was the Spanish indirect object clitic in the third person singular and plural forms (*le, les*). The indirect object clitic appears as a pronominal clitic element in double object constructions, as illustrated in the following sample sentences:

(1) a. *Juan le da un abrazo*
         Juan him/her-3RD.SING.DAT. give a hug
         'Juan gives him/her a hug'
     b. *Juan les envía una carta*
         Juan them-3RD.PL.DAT. send a letter
         'Juan sends them a letter'

Prior SLA research has shown that object clitics are challenging for L2 learners of Spanish to master (e.g., Leeman, 2003; Liceras et al., 1997; Ortega & Long, 1997; VanPatten, 1984). One of the main factors that contributes to their learning difficulty is their low level of perceptual and semantic or functional salience (e.g., Goldschneider & DeKeyser, 2001; Ortega & Long, 1997). From a perceptual perspective, clitics are not phonetically stressed in Spanish nor are they visually marked with a written accent. They also bear low relative sonority based on Laver's (1994) sonority hierarchy. Additionally, clitics hold a non-transparent relationship to their

pronominal referent and show low semantic weight (Ortega & Long, 1997). These and other properties make the indirect object clitic a particularly complex target structure in L2 Spanish.

Procedure

The study reported in this chapter is part of a larger research project examining task-based L2 development across different modality conditions (Zalbidea, 2018). Data for this study come from the second session of this larger project, where participants in the Writing and Speaking groups completed two computerized form-focused tasks, one of which targeted the indirect object clitic. Immediately after each focused task, participants filled out a post-task questionnaire regarding their perceptions of task demands. After completing a series of assessment tasks, a subset of participants in each group partook in individual stimulated recall interviews. The following section provides more detailed information about each of the three data sources considered in this study.

Materials

Participants completed the task individually on a computer. In the pre-task stage, all participants read a story in English that provided the necessary background information and instructions to complete the subsequent focused task. The story informs participants that a famous lab scientist has discovered a chemical substance that affords human superpowers and that, unfortunately, rumors about a potential robbery of the substance have begun to surface. The scientist has narrowed it down to two suspects, which are two research assistants that joined the lab soon after the discovery. The participant's job is to help the scientist obtain information about each suspect's routine through each of the focused tasks, acting as a detective. In the post-task stage, participants write a brief report for the scientist, indicating who they think is behind the rumored robbery.

Focused task
For each of the items in the computerized focused task, participants read a prompt related to the suspect's activities and then decided which of two possible events provided a logical follow-up to the prompt. Events were presented with a representative picture and a verbal phrase in the infinitive form. Participants in both groups provided their answer by creating a full sentence that incorporated the event they had selected, making any necessary minimal changes or additions. The Writing group typed their sentence into a text box, while the Speaking group said the sentence out loud into a microphone. After providing their response, both groups received

feedback in the form of targetlike model utterances that contained the indirect object clitic, as well as the picture that represented the logical event. The Writing group read the feedback in written format, whereas the Speaking group listened to the feedback in audio format. Learners were not informed that the focused task introduced them to the Spanish indirect object clitic, nor were they provided with any explicit information about the target form at any point during the task. The focused task contained 16 critical items (half targeting the singular and half targeting the plural 3rd person indirect object clitic) and 4 distractors.

Stimulated recall protocols

Stimulated recall protocols (Gass & Mackey, 2017) were used to collect data on participants' noticing behaviors during the focused task. Participants were shown screen-captured video or audio recordings along with the focused task files in order to enhance their memory for task performance. Following Gurzynski-Weiss and Baralt (2014), a stimulated recall protocol guide was used to conduct the interviews. For the first three to four items in the task, participants were asked to talk about their thoughts regarding (1) their event choice, (2) the sentence they had produced, and (3) the computerized feedback they had received.

Posttask questionnaire

Immediately after completing the focused task, participants responded to a post-task questionnaire that gathered their perceptions on a series of relevant task-related dimensions. The questionnaire included 9-point Likert scale items (e.g., 1 = *The task was very easy*; 9 = *The task was very difficult*). The questionnaire items, most of which had been adapted from Robinson (2001) and other prior studies (e.g., Baralt, 2013; Révész, Michel, & Gilabert, 2016; Sasayama, 2016), focused on the following task-related dimensions: (1) perceived mental effort, (2) task difficulty, (3) ratings of stress, (4) perceived task performance, (5) timing/rushedness, (6) task interest, (7) anxiety, (8) perceived linguistic difficulty in terms of input demands, and (9) perceived linguistic difficulty in terms of output demands.

Coding and scoring

Learners' introspective comments in the stimulated recall protocols were coded following a bottom-up approach. The following levels of L2 analysis were identified in the data: (1) *noticing*, (2) *searching*, and (3) *integrating* (see Zalbidea, 2018). A protocol was coded as evidencing *noticing* when participants expressed having noticed or referred to the target structure, its grammatical category, or its morphosyntactic environment. *Searching* was evidenced when participants claimed having looked for patterns, engaging in exemplar comparison, or noticing contingencies

in the input. Lastly, a protocol was coded as showing *integrating* when participants expressed having attached a meaning to the target structure. Integrating was found to be *unsuccessful* or *successful* based on whether learners' form-meaning connection was target-like or not.

Participants' performance on the clitic-focused task was examined for instances of form incorporation. They were awarded 1 possible point per response: half a point if a clitic form was supplied in their sentence, and half a point if the target structure was morphosyntactically accurate. Participants received a 0 score if a clitic form was not supplied. Finally, participants' ratings on the post-task questionnaires were averaged and compared at the group level for each task dimension.

Analyses

Group comparisons for the form incorporation and post-task questionnaire data were conducted using non-parametric Mann-Whitney *U*-tests. Cohen's *d* was calculated as an effect size measure and interpreted following Plonsky and Oswald (2014): .40 was deemed small, .70, medium; and 1.00, large.

Results

Stimulated recall protocols: Noticing

The first part of the research question asked about the extent to which task modality influences learners' noticing of L2 grammar. Introspective comments from 10 focal participants (*n* = 5 in each group) were examined. Participants were primarily selected on account of their language background (all were native English speakers) and non-exclusion status in the study (none reported looking up information about the target structures of the research project outside of the study).

Tables 2 and 3 present a summary of the coding for the stimulated recall protocols in the Writing and Speaking group, respectively. As shown in the tables, all focal participants were found to engage in all three major levels of L2 analysis – noticing, searching, and integrating – regardless of their task modality assignment. Notably, three participants in the Writing group and two in the Speaking group evidenced successful integration of form and meaning, with the remainder of focal participants in each group showing unsuccessful integration. Unsuccessful integration resulted from various forms of non-targetlike semantic integration, including associating the indirect object clitic with the subject noun phrase (Participant 3, Participant 4, Participant 9) and the direct object noun phrase (Participant 8, Participant 9), instead of with the indirect object noun phrase.

Table 2. Writing group: Coding of the stimulated recall protocols

Participant	Reported levels of analysis				Sample comments
	Notice	Search	Integrate		
			Unsuccessful	Successful	
P-1	✓	✓	–	✓	"It's like object, object pronouns or whatever it is in Spanish, where you put *le* or *les*" / "It was like 'they gave *them* their number'"/ "There had been *le* or *les* in the other sentences and it was like 'they were asking *him* for help'"
P-2	✓	✓	–	✓	"I noticed (…) the *le* (…) it was a little bit more complex than I first imagined" / "Instead of 'I ask for help', it's like 'I ask *you* for help; (…) is that the object, is it? I don't know. I forget what it's called, but like it's like the action is being done to"
P-3	✓	✓	✓	–	"I'd gotten in my head that *le* would go with *él* and *ella*, and *les* will go with *ellos*, so I think I made that mistake for a while until I got one wrong" / "Eventually I figured maybe I had to look further in the sentences" / "I thought it had something to do with possession, but I don't think that's right"
P-4	✓	✓	✓	–	"The *le* part confused me. 'Why there needs to be a *le*?'" / "If it's *ellas*, it's *le*; if it's *ellos*, it's *les*. Maybe?" / "I tried to get what before taught me about the *les* (…) I mean, I still didn't completely understand it"
P-5	✓	✓	–	✓	"I noticed that it was *le* you use to say 'give *to him*' / But then I learned that you had to put '*les* piden' once I saw [the feedback] / I think *le* means 'him' or 'her' or… 'for him' or 'for her', and then the *les* means 'them' or 'to them', 'for them', something like that"

Table 3. Speaking group: Coding of the stimulated recall protocols

Participant	Reported levels of analysis				Sample comments
	Notice	Search	Integrate		
			Unsuccessful	Successful	
P-6	✓	✓	–	✓	"I heard the extra word" / "Like if you're saying like 'asks for' or something like that, it sort of is like the people that you're asking I think… it like shows not subject but like the object of the verb I think, yeah"
P-7	✓	✓	–	✓	"Well I didn't know where the *le* came from" / "I started to realize that like you have to like put it in. I don't know (…) what these words are called, just like *le*" / "It's like who is doing it is how you conjugate the verb and then who's being affected by you doing that is that what that word is"
P-8	✓	✓	✓	–	"I noticed that they put like a *le* in front of the verb" / "I think it's cause they were exchanging phone numbers which is a plural object so that's why it was like *les* instead of *le*" / "I put the *le* in front of the verb this time because I kind of thought that like the object of the sentence was *ayuda*, which is a singular noun"
P-9	✓	✓	✓	–	"I was like 'What's the *le* and *les* before the verb?' I don't remember learning that" / "I'm not quite sure what that was, because it had nothing to do with the number of the subject" / "I thought that it was the gender like the subject, but then they weren't like paired" / "I thought maybe it was the object, (…) 'they asked themselves' or something, but I guess it would be reflexive"
P-10	✓	✓	✓	–	"They used like *le* and *les* (…) I didn't know how to use that" / "I thought that the *les* would go with plural subject. It didn't seem to always do that though" / "My next thought was that (…) it looked to the object / "It seemed that more often it was the opposite of what I thought"

In sum, examinations of the stimulated recall protocols do not appear to reveal major differences in noticing or further L2 analysis between the Writing and the Speaking groups. All focal participants in both groups engaged in the highest level of L2 analysis (i.e., integrating) during the focused task, above and beyond noticing of the clitic form. Instances of success in integrating form-meaning connections were also relatively similar across both modality groups.

Focused task: Form incorporation

The global research question of this study also asked about the extent to which task modality influenced learners' form incorporation of L2 grammar, that is, whether the noticed form was incorporated in subsequent performance. As can be observed in the descriptive statistics presented in Table 4, both groups evidenced a considerable amount of target form incorporation during task performance. Closer examination of the table reveals that participants in the Writing group supplied the clitic form at a greater rate than participants in the Speaking group, and that the morphosyntactic accuracy of the forms they incorporated into their own output was also superior.

The Mann-Whitney U-tests reported on Table 5 revealed that the Writing group significantly outperformed the Speaking group in overall form incorporation. Further group comparisons indicated that, whereas groups were not statistically different in terms of clitic form supply, the Writing group exhibited significantly higher accuracy rates in their clitic form incorporation than the Speaking group. The effect sizes of these significant differences were of small-to-medium magnitude.

Table 4. Descriptive statistics: Form incorporation by group

	Writing				Speaking			
	M (SD)	*Mdn*	Min.	Max.	*M (SD)*	*Mdn*	Min.	Max.
Overall incorporation	.55 (.22)	.55	.03	.91	.43 (.21)	.44	.00	.84
Form supply	.70 (.22)	.75	.06	.94	.61 (.25)	.63	.00	.94
Form accuracy	.53 (.24)	.54	.00	1.00	.38 (.23)	.36	.00	.93

Note. Writing, $n = 28$; Speaking, $n = 27$.

Table 5. Mann-Whitney U-tests comparing form incorporation in Writing and Speaking groups

	U	*z*	*p*	*d* [95% CI]
Overall incorporation	249.50	−2.17	.030*	−.58 [−1.13, −.03]
Form supply	273.50	−1.77	.076	−.40 [−.95, .15]
Form accuracy	221.00	−2.30	.021*	−.67 [−1.23, −.10]

* *Note p < .05*

Taken together, these results indicate that both the Writing and the Speaking modality conditions promoted comparable levels of clitic form incorporation during the task. However, participants in the Writing group incorporated the clitic form into their own output with superior morphosyntactic accuracy than participants in the Speaking group.

Post-task questionnaire: Perceptions of task demands

The final dimension targeted in our research question concerned learners' task demand perceptions. Descriptive statistics for the Writing and Speaking groups' post-task questionnaire ratings are shown on Table 6. As indicated by the distribution of means and medians, participants in both modality conditions provided relatively similar ratings across all nine task dimensions. Descriptive group differences were most apparent for ratings of task-induced stress and anxiety, both of which were greater in the Speaking group compared to the Writing group. However, the Mann-Whitney U-tests revealed no significant group differences in post-task ratings for any task dimension, as reported in Table 7.

Table 6. Descriptive statistics: Post-task questionnaire ratings by group

Task dimension	Writing				Speaking			
	M (SD)	*Mdn*	Min.	Max.	*M (SD)*	*Mdn*	Min.	Max.
Mental effort	5.57 (1.50)	6.00	2	8	5.65 (1.44)	5.50	3	9
Task difficulty	5.21 (1.55)	5.50	3	9	5.50 (1.36)	5.50	2	8
Stress	3.04 (1.53)	3.00	1	6	3.77 (1.80)	4.00	1	7
Task performance	4.25 (1.48)	4.00	1	7	4.62 (1.60)	5.00	2	8
Rushedness	2.25 (1.46)	2.00	1	6	2.58 (1.63)	2.00	1	6
Interest	4.57 (1.53)	5.00	1	6	4.85 (1.87)	5.00	2	9
Anxiety	2.43 (1.57)	2.00	1	7	3.00 (1.88)	2.50	1	7
Input difficulty	4.71 (1.90)	4.00	1	9	4.52 (1.64)	5.00	1	7
Output difficulty	4.82 (1.57)	5.00	2	7	4.71 (1.64)	5.00	1	7

Note. Writing, $n = 28$; Speaking, $n = 26$. Task performance was reverse-coded.

Table 7. Mann-Whitney *U*-tests comparing post-task ratings
in writing and speaking groups

Task dimension	*U*	*z*	*p*	*d* [95% CI]
Mental effort	358.50	−.10	.922	−.06 [−.60, .49]
Task difficulty	318.50	−.81	.421	−.20 [−.74, .35]
Stress	281.00	−1.46	.144	−.44 [−.99, .11]
Task performance	314.00	−.88	.378	−.24 [−.79, .31]
Rushedness	323.00	−.74	.460	−.21 [−.76, .34]
Interest	364.00	.00	1.000	−.16 [−.71, .39]
Anxiety	303.50	−1.08	.282	−.33 [−.88, .22]
Input difficulty	355.50	−.15	.881	.11 [−.44, .66]
Output difficulty	346.00	−.09	.928	.06 [−.49, .62]

In sum, the Writing and Speaking groups were highly similar in how they rated questionnaire items tapping their perceptions of the cognitive and performance demands of the focused task, among other relevant aspects. Whereas higher ratings of stress and anxiety were observed in the Speaking group compared to the Writing group, these differences were not statistically significant.

Discussion

The goal of this study was to investigate the extent to which task modality (writing vs. speaking) impacts learners' (a) noticing, (b) form incorporation, and (c) perceptions of task demands in an L2 grammar-focused task. To that end, three data sources were examined: (a) stimulated recall protocols were employed to explore learners' noticing behaviors, (b) learners' task performance was assessed for instances of form incorporation, and (c) subjective post-task questionnaires were administered to gauge learners' perceptions of task-generated demands. This section summarizes the main findings obtained from each of these data sources, identifies how they fit with relevant prior research, and discusses potential interpretations and implications of the study by taking all three information sources into consideration. The section concludes by assessing the relevance of the findings for current discussions of the language learning potential of L2 writing.

Concerning noticing, findings from the stimulated recall protocols revealed no major differences on the basis of task modality. All focal participants in both the Writing and Speaking groups reported having engaged in noticing of the indirect object clitic. Moreover, above and beyond noticing, analyses of the participants' introspective comments showed that all learners had engaged in searching as well as integrating, which was identified as the highest level of L2 analysis, pointing

to deep processing of the target form in both groups. Additionally, similar cases of both successful and unsuccessful form-meaning integration were present in both groups.

These findings are consistent with recent work by Ziegler (2017) and Gurzynski-Weiss and Baralt (2015), where task modality was not found to be a determining factor for rate of noticing as measured by stimulated recall protocols. Results do not seem to follow Yuksel and Inan (2014) or Lai and Zhao (2006), where some form of advantage was found for the written modality in terms of noticing of communication breakdowns. It should be noted, however, that these prior studies differ widely from the present study in that, in earlier research, participants completed unfocused interactive tasks in which they received feedback from a human interlocutor on a variety of targets (e.g., grammar, lexis). In this study, participants completed a computerized focused task where they received feedback in the form of target-like model utterances. As a result, the possible range of linguistic material to be noticed and compared was more controlled in the present study. Thus, it is likely that the design of the focused task employed here facilitated greater opportunities for learner-generated noticing and deeper processing among both learner groups, regardless of task modality.

In terms of form incorporation, initial analyses of learners' task performance revealed an advantage for the Writing group over the Speaking group in overall form incorporation. Subsequent analyses showed that groups evidenced comparable rates of clitic form incorporation, but that the Writing group outperformed the Speaking group in terms of morphosyntactic accuracy of the clitic form. This pattern of results indicates that the design of the focused task used in the present study was successful in promoting substantive form incorporation among participants regardless of the modality in which the output-input practice was conducted; nonetheless, producing the L2 and receiving model feedback in the written modality were shown to facilitate more accurate output modification.

These results do not follow earlier research by Gurzynski-Weiss and Baralt (2014), Lai and Zhao (2006), and Ziegler (2017), where no advantages were found in terms of modified output following recasts for the written (i.e., text-chat) modality over the oral (i.e., face-to-face) modality. As noted earlier, however, comparisons should be drawn with care because the task design in the present study was unlike in previous studies, and perhaps more crucially, form incorporation was operationalized differently. In particular, the computerized focused task used in the present study did not build opportunities for modified output following feedback into the task; rather, output modification was measured as the extent to which participants incorporated the indirect object clitic into their production throughout the completion of the task. Nonetheless, consistent with earlier conceptualizations of modified output, it was assumed that changes in learners' output in the present study resulted

from learners' detection of mismatches between their L2 production and relevant target-like input in the feedback.

Lastly, regarding perceptions of task demands, both the Writing and Speaking groups provided relatively similar ratings for the range of task dimensions considered in the study. While the largest descriptive group differences were observed for ratings of task-induced stress and anxiety, both of which were higher for the Speaking group, these were not statistically significant. This might suggest that focused task requirements (i.e., producing and processing sentence-level L2 output and input), which were the common denominator between groups, were most important in determining learners' task demand ratings in the study.

In general, these findings do not appear to align with prior research by Baralt (2013) and Cho (2018), where oral tasks were perceived as significantly more difficult than written tasks. Cho (2018) further reported that learners perceived themselves as more competent in completing written tasks; nonetheless, in the present study learners reported comparable levels of perceived ability to successfully perform written and oral versions of the focused task. Once again, it is worth highlighting the methodological differences across these studies in integrating findings with earlier work. In addition to the types of tasks that learners completed, one key difference is that participants in Baralt (2013) and Cho (2018) were both intermediate-level L2 learners, whereas participants in this study were beginner-level learners with limited working knowledge of the target structure. Crucially, the focused task in this study was designed to engage substantive cognitive resources with the aim of boosting opportunities for L2 development among lower proficiency learners (Izumi, 2002). Given that subjective task demands are derived from the interplay between the design of the pedagogic task and individual learners' L2 abilities (Bachman, 2002), it is likely that these task requirement and participant factors contributed to some of the similarities in task perceptions observed between modality groups in this study.

In sum, the study found advantages for the written modality over the oral modality in terms of form incorporation, whereas both groups were relatively similar with regards to noticing and subjective task demands. These results may be explained with reference to the modality-inherent characteristics of writing and speaking as well as the specific methodological aspects of this study. Thus, a plausible explanation for the finding that the written modality promoted more morphosyntactically accurate incorporation of the clitic form compared to the oral modality may be that the slower-paced, visual and more permanent nature of the written modality provided greater opportunities for participants to identify mismatches between their output and the model feedback (e.g., Gilabert et al., 2016). This lower-pressure environment (e.g., Manchón, 2014; Williams, 2012), in turn, may have allowed learners more time and support to successfully disentangle the

complex form-to-meaning relationships that characterize the indirect object clitic, subsequently leading to more accurate productions of the form in learner output. In contrast, the faster and transient nature of oral input appears to have placed a greater load on learners' attentional resources (e.g., Gilabert et al., 2016), possibly leading to less efficient output-input comparisons online, and ultimately resulting in less accurate incorporations of the target form.

The group similarities attested for noticing may be explained with reference to the nature of the focused task as well as the specific measure of noticing employed in the present study. As noted earlier, the focused task required learners to produce sentences, after which they received feedback in the form of model input, and was purposefully structured to boost opportunities for noticing and further processing among beginner learners (Izumi, 2002). It is thus likely that the focused and more controlled nature of the task effectively provided opportunities for deep linguistic processing in both modality groups, leading to high levels of reported L2 analysis among all focal participants. Alternatively, given the modality differences attested for form incorporation accuracy, it is also possible that task modality did promote certain differences in learners' noticing behaviors, but that these differences were not captured in the stimulated recall data due to the offline nature of this noticing measure. For instance, it may be that certain group differences existed in the timing or amount in which learners engaged in each level of L2 analysis, which might have contributed to the quantitative differences found in the form incorporation data.

Similarly, the lack of major differences observed in learners' ratings of task demands may also be explained with reference to the cognitive requirements of the focused task. It is possible that the relatively high linguistic and content de-mands of the focused task, which were common across both modality conditions, were the primary factors driving beginner learners' ratings in the task perception questionnaire, more so than output/input modality. Nonetheless, the descriptive differences found for ratings of task-induced stress and anxiety, which were higher in the oral task, are not unexpected under the assumption that learners experience greater attentional pressure to produce output and process input online in the oral modality compared to the written modality (e.g., Gilabert et al., 2016; Vasylets et al., 2017; Williams, 2012; Zalbidea, 2017).

In short, these results allow us to shed light on the L2 learning potential of writing compared to speaking. In particular, the fact that advantages were observed in the Writing group over the Speaking group in terms of form incorporation of the clitic form during the focused task suggests that the written and oral modal-ities indeed may differ in their affordances for L2 grammar development. More specifically, the finding that output-input practice in the written modality led to more morphosyntactically accurate form incorporation, echoing the establish-ment of more target-like form-meaning mappings among learners in that group,

is indicative of a potential advantage for restructuring (McLaughlin, 1990) in the written modality, as suggested by Skehan (1996). These advantages are derived from the conditions for output and input processing that characterize each modality. Thus, it is reasonable to assume that the visual quality of written language as well as its slower and more self-regulated nature can promote conditions that are most propitious for rapid and successful learner-generated focus on form.

Conclusions and limitations

Findings from the present study provide some evidence suggesting that the L2 learning potential of focused tasks can be enhanced in the written modality compared to the oral modality, consistent with Gilabert et al.'s (2016) postulations. Although both written and oral tasks promoted high levels of L2 analysis, written tasks gave way to more accurate incorporations of the target form throughout the output-based form-focused task. Lastly, learners in both modality conditions reported comparable subjective task demands, although ratings of task-induced stress and anxiety were descriptively higher in the oral modality. These findings should be interpreted in light of a number of methodological limitations that must be acknowledged. First, due to logistic constraints, only a subset of participants was considered for the stimulated recall data, which restricts generalizability at the group level. Additionally, stimulated recall protocols are limited with regards to the potential issues of veridicality and memory decay (Leow, 2015), so learners' introspective comments should be interpreted with these methodological idiosyncrasies in mind. Lastly, the study reported in this chapter did not consider L2 development data, and therefore, the extent to which the indicators of L2 learning potential considered here actually promote differential L2 gains remains to be addressed.

Notwithstanding the aforementioned limitations, the present study highlights the usefulness of adopting mixed-methods approaches that consider multiple data sources in order to investigate the L2 learning potential of L2 writing. Prior studies have predominantly focused on examining learner introspective comments as a primary (and sometimes the only) source of information into a task's potential to promote L2 outcomes. Such an approach would have provided only a partial view into the role of task modality in this study. Instead, combining data sources using a mixed-methods approach proved most informative in assessing the fuller effects of task modality and in making more robust and reliable insights possible (e.g., Riazi, 2016). The stimulated recall data provided rich, qualitative information into learners' noticing behaviors during task completion, revealing high levels of L2 analysis and evidencing the use of various successful and unsuccessful strategies

among individual learners. In turn, learners' task performance offered relevant quantitative data to examine whether, beyond noticing, participants actually incorporated the target form into their own output and whether they did so accurately, in ways that were observable at the group level. Finally, the post-task questionnaire provided valuable quantitative information on learners' perceptions, particularly as they pertained to dimensions of processing and performance demands, among other aspects, allowing us to better characterize learner experiences during task completion in both modality conditions.

In sum, this study has provided new insights into the language learning potential of L2 writing compared to speaking, showing that focused tasks involving output-input cycles in the written modality can provide favorable conditions for promoting learner-generated focus on L2 grammar forms, as evidenced by more target-like integration of grammatical constructions into learners' output. These findings add to the increasing evidence pointing to modality as an influential task design factor in L2 research (e.g., Vasylets et al., 2017; Zalbidea, 2017), and highlight the empirical value of considering multiple qualitative and quantitative sources of data on L2 processes and products to better understand and conceptualize the potential of L2 writing as a catalyst to L2 development.

References

Adams, R. (2006). L2 tasks and orientation to form. *International Journal of Applied Linguistics*, 152, 7–33. https://doi.org/10.2143/ITL.152.0.2017861

Adams, R., & Ross-Feldman, L. (2008). Does writing influence learner attention to form? In D. Belcher & A. Hirvela (Eds.), *The oral-literate connection: Perspectives on L2 speaking, writing, and other media interactions* (pp. 243–266). Ann Arbor, MI: University of Michigan Press.

Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language Testing*, 19, 453–476. https://doi.org/10.1191/0265532202lt240oa

Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition*, 35, 689–725. https://doi.org/10.1017/S0272263113000429

Cho, M. (2018). Task complexity and modality: Exploring learners' experience from the perspective of flow. *Modern Language Journal*, 102, 162–180. https://doi.org/10.1111/modl.12460

Cleland, A. A., & Pickering, M. J. (2006). Do writing and speaking employ the same syntactic representations? *Journal of Memory and Language*, 54, 185–198. https://doi.org/10.1016/j.jml.2005.10.003

Gass, S. M., & Mackey, A. (2017). *Stimulated recall methodology in applied linguistics and L2 research* (2nd ed.). London: Routledge.

Gilabert, R., Manchón, R. M., & Vasylets, L. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, 117–135. https://doi.org/10.1017/S0267190515000112

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A Meta-analysis of multiple determinants. *Language Learning*, 51, 1–50. https://doi.org/10.1111/1467-9922.00147

Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in FTF and CMC modes. *Studies in Second Language Acquisition*, 36, 1–37. https://doi.org/10.1017/S0272263113000363

Gurzynski-Weiss, L., & Baralt, M. (2015). Does type of modified output correspond to learner noticing of feedback? A closer look in face-to-face and computer-mediated task-based interaction. *Applied Psycholinguistics*, 36, 1393–1420. https://doi.org/10.1017/S0142716414000320

Izumi, S. (2002). Output, input enhancement, and the Noticing Hypothesis. *Studies in Second Language Acquisition*, 24, 541–577. https://doi.org/10.1017/S0272263102004023

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62, 439–472.
https://doi.org/10.1111/j.1467-9922.2012.00695.x

Kuiken, F., & Vedder, I. (2012). Speaking and writing tasks and their effects on second language performance. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 364–377). London: Routledge.

Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology*, 10, 102–120.

Laver, J. (1994). *Principles of phonetics*. Cambridge, MA: CUP. https://doi.org/10.1017/CBO9781139166621

Leeman, J. (2003). Recasts and second language development. *Studies in Second Language Acquisition*, 25, 37–63. https://doi.org/10.1017/S0272263103000020

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge. https://doi.org/10.4324/9781315887074

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT press.

Liceras, J., Maxwell, D., Laguardia, B., Fernández, Z., Fernández, R., & Díaz, L. (1997). A longitudinal study of Spanish non-native grammars: Beyond parameters. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish, Vol. 1: Developing grammars* (pp. 99–132). Somerville, MA: Cascadilla Press.

Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 27–53). Amsterdam: John Benjamins.

McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32, 207–224. https://doi.org/10.1016/j.system.2004.01.003

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition*, 27, 79–103.
https://doi.org/10.1017/S0272263105050047

McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11, 113–128.
https://doi.org/10.1093/applin/11.2.113

Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness*, 18, 384–402. https://doi.org/10.1080/09658410903197256

Ortega, L. (2012). Epilogue: Exploring L2 writing–SLA interfaces. *Journal of Second Language Writing*, 21, 404–415. https://doi.org/10.1016/j.jslw.2012.09.002

Ortega, L., & Long, M. H. (1997). The effects of models and recasts on the acquisition of object topicalization and adverb placement in L2 Spanish. *Spanish Applied Linguistics*, 1, 65–86.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912.   https://doi.org/10.1111/lang.12079

Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition*, 38, 703–737.   https://doi.org/10.1017/S0272263115000339

Riazi, A. M. (2016). Innovative mixed-methods research: Moving beyond design technicalities to epistemological and methodological realizations. *Applied Linguistics*, 37, 33–49.
https://doi.org/10.1093/applin/amv064

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57.
https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2003). The Cognition Hypothesis, task design, and adult task-based language learning. *Second Language Studies*, 21(Spring), 45–105. University of Hawaii, Dept. of Second Language Studies.

Rouhshad, A., Wigglesworth, G., & Storch, N. (2016). The nature of negotiations in face-to-face versus computer-mediated communication in pair interactions. *Language Teaching Research*, 20, 514–534.   https://doi.org/10.1177/1362168815584455

Sagarra, N., & Abbuhl, R. (2013). Optimizing the noticing of recasts via computer-delivered feedback: Evidence that oral input enhancement and working memory help second language learning. *Modern Language Journal*, 97, 196–216.
https://doi.org/10.1111/j.1540-4781.2013.01427.x

Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of cognitive task complexity. *Modern Language Journal*, 100, 231–254.   https://doi.org/10.1111/modl.12313

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, UK: CUP.   https://doi.org/10.1017/CBO9781139524780.003

Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32, 203–234.
https://doi.org/10.1017/S0272263109990507

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.   https://doi.org/10.1093/applin/17.1.38

VanPatten, B. (1984). Learners' comprehension of clitic pronouns: More evidence for a word order strategy. *Hispanic Linguistics*, 1, 57–68.

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67, 394–430.
https://doi.org/10.1111/lang.12228

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331.   https://doi.org/10.1016/j.jslw.2012.09.007

Yilmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, 62, 1134–1169.
https://doi.org/10.1111/j.1467-9922.2012.00726.x

Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research*, 15, 457–477.
https://doi.org/10.1177/1362168811412873

Yuksel, D., & Inan, B. (2014). The effects of communication mode on negotiation of meaning and its noticing. *ReCALL*, 26, 333–354.   https://doi.org/10.1017/S0958344014000147

Zalbidea, J. (2017). 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *Modern Language Journal*, 101, 335–352. https://doi.org/10.1111/modl.12389

Zalbidea, J. (2018). Output, task modality, and target form salience in learner noticing of form and L2 development (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Ziegler, N. (2017). Task modality, noticing, and the contingency of recasts: Insights on salience from multiple modalities. In S. M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 269–290). New York, NY: Routledge. https://doi.org/10.4324/9781315399027-14

Processing output during individual L2 writing tasks

An exploration of depth of processing and the effects of proficiency

Sonia López-Serrano[1,2], Julio Roca de Larios[2] and Rosa M. Manchón[2]
[1]Pompeu Fabra University / [2]University of Murcia

The present study explored the different levels of processing as well as the strategic orientation involved in the language-related episodes identified in the think-aloud protocols produced by 21 EFL writers (divided into three proficiency groups: pre-intermediate, intermediate and advanced) during the individual completion of an argumentative writing task. Results confirmed previous predictions on (i) the manner in which engaging in L2 writing may foster deep levels of language processing, and (ii) the proficiency-dependency of such processing. In particular, advanced learners engaged in deeper levels of processing with an upgrading strategic orientation, which involved complex manipulation and evaluation of L2 forms and structures. These findings are discussed through the lens of the language learning potential of individual L2 writing tasks.

Introduction

As discussed in the Introduction to this volume, the study of writing as a site for language learning is a recently opened research avenue in instructed second language acquisition (ISLA) studies. Yet, arguments in support of a connection between writing and language learning go back to the early 1990s and, more precisely, to Cumming's (1990) pioneering paper in which he made a strong case for the important language learning gains that could potentially derive from writing events characterised by an intense linguistic meaning-making activity. Importantly, despite the abundant theoretical and empirical work on the connection between writing and language learning that has been produced since, Cumming's original proposal for moving research forward has been almost ignored, especially his call for descriptive,

exploratory work on those individual composing processes anticipated to "have potential for learning of the language" (1990, p. 484), as a preliminary step before embarking in more controlled studies.

The research reported in this chapter is an attempt to advance in this direction. We aim to provide the kind of descriptive data that could be useful before conducting experimental and longitudinal studies (see also Chapter 14, this volume). In the framing of the study that follows, we synthesize relevant theoretical positions on the connection between writing and language learning with the ultimate aim of highlighting the relevance of prioritizing a focus on writing processes in individual writing conditions, the type of writing under the spotlight in our own study. We will then review the available (albeit scant) empirical research on writing processes undertaken from a writing-to-learn language perspective.

Theoretical and empirical background

Writing as a site for language learning: The relevance of studying individual writing

A key element in Cumming's (1990) pioneering work on writing and language learning was the concept of "comprehensible output" (Swain, 1985), which allowed him to formulate the psycholinguistic rationale for the language learning potential of composition writing in the following way (Cumming, 1990, p. 483):

> Composing might function broadly as a psycholinguistic output condition wherein learners analyze and consolidate second language knowledge that they have previously (but not fully) acquired […]. Composition writing elicits an attention to form-meaning relations that may prompt learners to refine their linguistic expression – and hence their control over their linguistic knowledge – so that it is more accurately representative of their thoughts and of standard usage.

Cumming (1990) additionally posited that any potential language learning gains that may derive from the "form-meaning" processes mentioned in the quotation above would be dependent on a combination of learner-internal and external conditions, such as availability of second language (L2) knowledge, personal significance of writing, and availability of time. Following Cumming's proposal, Ortega (2009, p. 62) suggests that L2 writers will engage in the kind of decision-making processes of concurrent reflection on meaning and form that may foster language learning only if they push themselves "to express their intended meaning more precisely or if the nature of what they are trying to do with words (…) is demanding, cognitively and linguistically". In this sense, composition writing tasks require

learners to engage in cognitively demanding processes at both linguistic and strategic levels because of two main reasons: First, the intense meaning-making activity that characterises complex forms of writing (see Byrnes & Manchón, 2014) and, second, the need to make strategic decisions on the distribution of attentional resources to fulfil tasks demands and goals set (Manchón & Roca de Larios, 2007).

Importantly, writing goals are rarely instantly achieved, which means writers frequently need to engage in solving problems of an *ill-defined* nature. These are problems that do not have a clear path that leads to a solution, have multiple solutions or even no solution at all, and are thus considered to require the use of higher-level critical thinking and the application of different skills, strategies, and approaches. The writing problems and goals involved in the production of connected, contextualized, coherent, and appropriate texts are claimed to belong to this category. Accordingly, it is mainly through the application of their available L2 knowledge and the activation of a full range of problem-solving strategies that learners may have the opportunity to reflect on linguistic concerns and, as a result, make the most of the language learning potential involved in writing (Manchón & Roca de Larios, 2007). As suggested above, this type of reflection would be similar to what Swain characterized as *pushed output*: Output that extends the linguistic repertoire of the learners as they attempt to express their intended meaning "precisely, coherently and appropriately" (1985, p. 248). It is also the kind of output that is claimed to be the most beneficial for the consolidation and expansion of learners' L2 knowledge (Manchón & Williams, 2016).

Following from the above, writing tasks may be particularly beneficial in terms of vocabulary acquisition and consolidation (see Schmitt, Chapter 15, this volume), since the need for words learners experience when composing a text may lead them to conduct lexical searches that may help them restructure and expand their L2 lexicon. This advantage of writing would be supported by the *Involvement Load Hypothesis* (ILH), which was propounded by Laufer and Hulstijn (2001) as a way of conceptualizing vocabulary learning, and is characterized by three components. The first component, *need,* is motivational in nature: When learners want to express their intended message but they do not know "the appropriate word for it" (2001, p. 14). In turn, *search* would refer to the attempt to find the word needed to express the intended message, and *evaluation* would involve processes of comparison between different candidates to check their adequacy or suitability to the context. Composition tasks are considered to have a high involvement load, as the three components are strongly activated throughout the writing process.

Despite these well-established arguments, and also despite Cumming's original recommendations for a focus on writing processes in individual composing tasks as a way of shedding light on how and why challenging writing may be conducive to

language learning, very little scholarly attention has been paid to the investigation of metalinguistic thinking processes implemented during individual writing tasks or to their language learning potential. The scant empirical research on the issue is reviewed next.

Writing and language learning opportunities in individual writing conditions: The empirical evidence

Two studies conducted in the early nineties (Cumming, 1990; Swain & Lapkin, 1995) constituted pioneering attempts to look into writing processes in individual writing conditions from the perspective of their potential for language learning.

Using non-metacognitive think-aloud protocols, Cumming (1990) analysed the *decision-making episodes* in which 23 francophone learners of English (at two proficiency levels, intermediate and advanced) attended to form and meaning concurrently while writing two texts in the L2. These episodes were found to include three types of cognitive activity: (1) searching out and assessing improved phrasing; (2) comparing cross-linguistic equivalents; and (3) reasoning about linguistic choices. In addition, when trying to find and evaluate words and phrases to convey their intended meaning, the participants engaged in *more extensive* and *less extensive searches*. The first type frequently consisted of lengthy reflections which involved a great deal of "mental effort" on the part of the writer, whereas the second kind of episodes usually were simple revisions that involved less extensive cognitive activity. Although less extensive searches could contribute to the automatization and reinforcement of prior knowledge, Cumming hypothesized that more extensive searches, as they involved deliberate conscious reflection on language, could be more favourable for refining and consolidating L2 knowledge and use.

Swain and Lapkin (1995) conducted a partial replication of Cumming's study in which they explored nine adolescent French immersion students' linguistic reflections while composing. These reflections were captured through think-aloud protocols and operationalized in terms of *language related episodes* (LREs), which focused on both form and meaning, or just on form. The participants engaged in two different degrees of reflection on form that were labelled as *simple inspection* and *complex thinking*. However, no further elaborations or examples of these categories were provided.

It should be noted that, whereas individual writing processes were not further explored in relation to their language learning potential for more than twenty years, a fruitful strand of research began to pay attention to the opportunities for language learning that may derive from students' engagement in collaborative writing tasks (see Storch, 2016). These opportunities have been found to arise either

through negotiations of form and provisions of feedback in studies that follow interactionist views of acquisition, or through scaffolding and co-construction of knowledge in socioculturally-oriented studies. As was the case with the two studies on individual writing mentioned above, this research has also shown that learners' LREs present different levels of engagement in language reflection although current taxonomies have often failed to capture the nature of the linguistic reflection involved. Therefore, these conceptualizations are not directly applicable to the study of individual writing processes.

Renewed interest in language reflection during individual written production tasks has re-emerged in some recent studies in which the concept of depth of processing (DoP) is made central to the analysis. DoP has been defined as "the relative amount of cognitive effort, level of analysis, and elaboration of intake, together with the usage of prior knowledge, hypothesis testing, and rule formation employed in decoding and encoding some grammatical or lexical item in the input[1]" (Leow, 2015, p. 204). According to Leow (2015), the notion of DoP would be close to that of *degrees of elaboration* which, as developed by Craik (2002), refers to the ways in which the processing of an item can be enriched to facilitate its storage in long-term memory and subsequent recall. Such enrichment entails adding more information about the new item to its memory trace by means of *within-item elaboration* strategies (encoding different characteristics of the item, e.g. physical, phonological, semantic, etc.) or *between-item elaboration* strategies (ways of creating connections with previously stored knowledge.). This conceptualization of DoP is closely related to the strategic behaviour L2 writers activate to solve their problems during composition writing. Consequently, the relevance of looking into problem-solving in writing through the lens of DoP is clearly warranted.

Two recent studies (Bergsleithner, 2019; López-Serrano, Roca de Larios, & Manchón, 2019) have explored the relationship between DoP and written output. Bergsleithner looked at (i) the role played by prior knowledge (old vs. recent) in the levels of processing (DoP) involved in the writing of a text; and (ii) the relationship between DoP and subsequent performance in post-tests. Two groups of Brazilian EFL university students with different degrees of old prior knowledge (OPK) and recent prior knowledge (RPK) of English irregular past tense forms were provided with a list of 10 regular and 10 irregular verbs and asked to complete a written narrative using each verb once. This was done while thinking aloud non-metacognitively. The data showed that the RPK group processed the irregular verbs in the past significantly more at a medium level and significantly less at a

---

1. Even though this definition refers to the processing of input, DoP is also considered to play a role in the generation of output.

low level than the OPK. That is, they "needed more cognitive effort to process deeper the target forms since their previous knowledge was not yet automatized" (Bergsleithner, 2019, p. 111). Additionally, higher levels of processing correlated with superior performance in the post-tests, providing additional support to the claim (e.g. Shing & Brod, 2016) that previous knowledge may benefit L2 learning, particularly by helping learners make associations within their existing knowledge and improve their memory processes (encoding, consolidation, and retrieval) in order to acquire new knowledge. It may be concluded that, when completing controlled tasks, the ability to activate appropriate knowledge plays a role in the knowledge processing stage. It remains to be seen to what extent this is also the case with more complex writing tasks.

López-Serrano et al. (2019) used Leow's (2015) conceptualization and descriptors of DoP as a basis to create a coding system that captured the specific nature of output processing during individual writing. They analysed the language-related episodes present in the non-metacognitive think-aloud protocols produced by 21 EFL learners during the completion of an argumentative writing task. Every LRE was coded according to five criteria: (1) length of the LRE and pausing behaviour; (2) number of alternatives generated and assessed; (3) analysis and manipulation of different levels of linguistic representation; (4) amount and variety of strategies deployed to solve the LRE; and (5) use of metalanguage. Three levels of processing were identified: One in which participants reprocessed their output through the almost immediate application of knowledge (non problem solving low), and two levels that involved different degrees of activation of problem-solving strategies and the formulation and assessment of alternative language forms or syntactic structures (problem solving medium and problem solving high). The study provides a psycholinguistically-motivated coding system of writing processes in individual composition tasks potentially conducive to language learning, hence attempting to contribute to advancing research agendas along the directions envisaged in Cumming's (1990) study.

The present study, based on the same data as López Serrano et al. (2019), attempted to go one step further. With the aim of shedding light on one of the learner-internal conditions mentioned by Cumming as crucial to the learning potential of writing, namely, availability of L2 knowledge, we adopted the categories in López-Serrano et al. (2019) to look into the potential proficiency-dependency of language reflection activity while writing. The motivation behind this decision is elaborated next.

The role of L2 proficiency in linguistic processing

The construct of L2 proficiency results from interactions of multiple components, such as linguistic competence, metalinguistic awareness, and the abilities to read, write, listen and speak appropriately in the L2 (Hulstijn, 2011). From an information-processing perspective, L2 proficiency is purported to encompass two main dimensions: (1) the L2 writer's lexical and syntactic repertoire; and (2) the availability of lexical units and syntactic structures in the form of ready-made procedures that are differentially accessed and retrieved, regardless of their structural complexity, by means of "chunking" (Miller, 1956) or "proceduralization" (Anderson, 1982). These dimensions are thus closely related to the notions of "prior knowledge" and "availability of L2 knowledge", which are claimed to have an impact on DoP (Leow, 2015) and on the language learning potential of writing tasks (Cumming, 1990). Although Bergsleithner (2019) explored the effects of old and recent previous knowledge of a specific grammatical structure on DoP of two groups otherwise similar in terms of their L2 proficiency, to our knowledge, no previous study has explored how different levels of general L2 proficiency may affect learners' DoP during L2 writing production.

L2 proficiency and depth of processing

Research on input processing has shed light on the effects of L2 proficiency on learners' DoP. Calderón (2013) investigated the depth of aural input processing of the subjunctive shown by 24 low and intermediate proficiency university students of Spanish as a foreign language. Intermediate learners engaged in significantly lower DoP than the low proficiency group, although they achieved higher levels of awareness. Based on these data, Leow (2015) suggested that, in input processing, "once awareness at the level of understanding is reached, high levels of depth of processing are not only unnecessary but also infrequent" (p. 221). However, this effect of L2 proficiency may not be directly transferable to output processing during individual composition writing given that in this condition learners initiate their own episodes of attention to form spontaneously rather than being prompted to do so by any form of external input.

L2 proficiency and learners' strategic orientation of problem-solving behaviour

Previous research on L2 writing processes has shown that L2 proficiency affects writers' goals and strategic orientations. In their analyses of EFL learners' writing behaviours, scholars at the University of Murcia (e.g., Manchón, Roca de Larios, & Murphy, 2009; Roca de Larios, Manchón, Murphy, & Marín, 2008) found that

their participants' problem-solving behaviour was guided by two purposes: compensatory and upgrading. A *compensatory* orientation results from writers' lack of access to the linguistic knowledge required to express their intended meaning, either because it has not yet been acquired or because it has not been fully automatized. In contrast, in *upgrading* episodes the behaviour is triggered by an effort to improve the expression of one's intended meaning. These studies also found that their participants, who had three different proficiency levels, activated strategies that were guided by both orientations but that the temporal distribution of these orientations varied depending on their L2 proficiency: Generally, advanced learners were found to spend nine times more time on upgrading concerns than on compensatory ones, while their lower level counterparts spent twice as much time on compensatory concerns than on upgrading ones.

Although, as explained above, writers' goals have been claimed to play an important role in learning through writing, no study has investigated how these goals may affect learners' engagement in different degrees of DoP, or how such engagement may be mediated by L2 proficiency (see Galbraith & Vedder, 2019). It could be hypothesized that higher proficiency learners will not require deep levels of processing to complete their L2 texts since, as proficiency develops, writers gain greater control of the L2 and fluency increases. However, it may also be speculated that with the gradual automatization of text generation processes, mental capacity is freed, and this may allow advanced writers to direct their attention to higher-order concerns and carefully reflect on their upgrading goals, especially if the task at hand is open-ended and meaning-oriented and presents high cognitive demands.

The present study

Research questions

Taking into account Cumming's (1990) original suggestions regarding the way in which the connection between writing and language learning ought to be approached from an empirical perspective, and considering also the more recent theoretical assumptions presented above on the role of DoP in bringing about language learning through output production, the following research questions guided the present study:

1. Does the depth of processing observed in the LREs produced by EFL learners vary as a function of L2 proficiency?
2. Does the strategic orientation of the problem-solving behaviour found at different depths of processing vary as a function of L2 proficiency?

Method

Participants

The participants in this study were 21 Spanish learners of English as a foreign language (EFL) in three different year groups in the Spanish educational system with 6, 9, or 12 years of previous instruction in English. They comprised three proficiency groups (pre-intermediate, intermediate and advanced), according to their scores in the Oxford Placement Test (OPT) (Allan, 1995). The Pre-intermediate group (P-INT) consisted of seven high school students (two males and five females) aged 16–17. Their scores in the OPT ranged from 100 to 108 (B1 in the CEFR). The Intermediate (INT) group was composed of seven female university students, aged 19–20, who were in the third year of a degree in Education and whose scores were between 140 and 157 (B2 in the CEFR). Finally, the Advanced group (ADV) consisted of one male and six female students aged 23–24. They had a score of between 174 and 190 in the OPT (C1 in the CEFR), and were recent graduates in a 5-year degree in English. An ANOVA test conducted on the three sets of scores showed statistically significant differences across the three groups $[F (2, 18) = 397.227, p < 0. 000]$.

Task and data collection procedures

Our participants were asked to complete the following argumentative writing task (prompt provided in English):

> Success in education is influenced more by the student's home life and training as a child than by the quality and effectiveness of the educational programme. Do you agree or disagree?. (Raimes, 1987)

The participants were given an hour to complete the task under think aloud conditions (using English, Spanish or both). They did so without access to dictionaries or any other type of external help as one of the purposes of the project was to investigate what participants could do without the aid of additional resources.[2]

Data coding and analysis procedures

Identification of LREs
The language-related episodes produced by the participants were identified in the TA protocols by two coders following Swain and Lapkin's (1995, p. 378) definition:

---

2.   We are therefore aware that these experimental conditions may not be extrapolated to cases in which L2 writers compose having access to external resources.

> any segment of the protocol in which a learner either (1) spoke about a language problem he/she encountered while writing and solved it either correctly or incorrectly [or left it unresolved] or (2) simply solved it without having explicitly identified it as a problem.

Every LRE was isolated and subjected to an in-depth coding of the process followed by the participant from the initial state in which a language concern was identified to the time when it was solved (or left unresolved). This entailed coding five elements within each episode: (1) the language forms produced; (2) every strategy applied; (3) the knowledge sources verbalized; (4) the explicit technical metalanguage used (if any); and (5) the changes introduced in the written text. All the episodes were coded first in terms of their DoP and then those that showed problem-solving behaviour were coded for their strategic orientation. Both dimensions are described and illustrated below.

Depth of processing

Data were analysed according to the three levels of processing identified by López-Serrano et al. (2019): non-problem solving low (NPS-L), problem solving medium (PS-M), and problem solving high (PS-H). The three levels are described in Table 1.

Table 1. Levels of processing

DoP		Description
Non problem solving low	Episodes that involve the (almost) immediate application of L2 (or L1) knowledge to solve the language concern at hand.	Brief episodes that focus on one or two forms that belong to one level of linguistic representation.
Problem solving medium	Episodes that involve the conscious activation of a number of problem-solving strategies to find a solution for the language concern at hand.	Episodes that include the production and assessment of around three L2 forms or the reformulation of a syntactic structure. They sometimes involve attention to different levels of representation.
Problem solving high		Long episodes that involve i. the generation and evaluation of a series of forms that frequently belong to multiple levels of linguistic representation, and ii. the application of a variety of problem-solving strategies, including resorting to technical metalanguage.

Strategic orientation

Those episodes that evidenced problem-solving behaviour were further analysed in terms of how strategies were orchestrated from the initial state to the final state of the LRE. Episodes were labelled as *compensatory* when the strategies deployed aimed at making up for a lack of L2 knowledge or at solving difficulties in accessing their linguistic resources. Those episodes whose aim was to refine and improve an intended message already available in the L2 by looking for a better match between their communicative intention and the language used to express it, were coded as *upgrading*. Finally, our data showed some cases in which no clear orientation was shown, which were coded as *no orientation*.

Examples[3]

Example (1) shows a non-problem-solving LRE in which the participant realized that the previously written form "I am agree" was incorrect and deleted the word "am" to turn it into the correct option "I agree". These alternatives were uttered at a fast pace and no pauses or other evidences of further reflection appeared in the episode.

(1) NPS-L
     I am agree mmm I am agree no I am agree no es I agree I agree   (INT6-E34)

Example (2) illustrates a problem-solving medium compensatory LRE in which the learner intended to translate the expression "se van de su casa" (move away from home). As can be seen in the protocol, the translation was conditioned by the need to use the verb "go", which was immediately retrieved as the equivalent to the Spanish "ir". From this point, the participant generated combinations of "go" with the prepositions "out" and "to", and reread and backtranslated his text to try to find a correct option. However, he finished the episode with the incorrect translation "go the their house".

(2) PS-M compensatory
     there are children que se van de su casa go go out (3) se van de su casa (5) go to de to *there are children también hay chicos hay chicos* que se van de su casa (3) que se van de su casa ¿cómo lo pongo? *También hay chicos* que se van de su casa that go ir **that go**(4) *que se van that go* ir de casa *that go* **the their house**
     (P-INT4–E49)

---

3.   The following conventions are used: All words in normal font are TA talk by the writer; segments in **bold** indicate text actually written down; *italics* indicate re-readings and repetitions; numbers in parentheses indicate pause duration in seconds. Each example has been assigned a code that refers to the participant that has produced the LRE (e.g. ADV3) and the number of that particular LRE (E4).

Finally, in Example (3) we can see a PS-H upgrading LRE in which the participant generated and evaluated eight alternative syntactic structures (and four verb phrases) in the L2 until she found a clause that successfully conveyed her intended message. In this case, her verbalizations show that she wanted to construct a clause that (1) accurately represented her reality and that of her sisters, and that was (2) linguistically sophisticated. Evidence of this appears in her explicit evaluations, such as considering that the structure "our academic success has been different" was "too simple" and "not original" at all.

(3)   PS-H upgrading
we have even been to the same school yet all our success (3) yet there has eso no no es lo que es que no sé cómo seguirlo ¿qué es lo que quiero poner? sin embargo *we have even been to the same school* I wanna say yet our academic success has been different ¡qué graciosa! ¡qué original! ¡qué básico! (5) yet we haven't been academically successful (3) we have not the same degree (7) yet there has always been a different vamos a ver voy a leer un poco *they have always encouraged us to read and study and we more or less have the same cultural background we've even been to the same school* **yet** (4) my youngest sister's (9) academic results no *yet* my youngest **my youngest sister** has usually (4) has always porque es always (5) **has always** (6) *has always* found difficulty in studying (3) *has always* had difficulty has (3) *has always* had lower academic *has always* been less successful eso es *has always* **been less successful** *has always been less successful* than the other (4) than my sister and me **than my other sister and me**   (ADV1-E35)

Results

Research question 1

Our first research question asked whether DoP varied as a function of L2 proficiency. Our data showed that writers at all proficiency levels engaged in reflection on language at the three DoP levels identified in our coding system. Importantly, even though LREs that involved high levels of processing were the least common for all participants, their number increased with proficiency: Advanced participants produced three times more PS-H episodes ($M = 8.57$) than their Pre-intermediate counterparts ($M=2.71$) (see Table 2).

Table 2. Depth of processing across proficiency levels

	P-INT			INT			ADV		
	N	M	SD	N	M	SD	N	M	SD
NPS-L	135	19.29	9.43	175	25.00	6.22	152	21.71	13.66
PS-M	122	17.43	8.66	155	22.14	6.72	182	26.00	9.76
PS-H	19	2.71	2.06	44	6.29	3.45	60	8.57	6.73
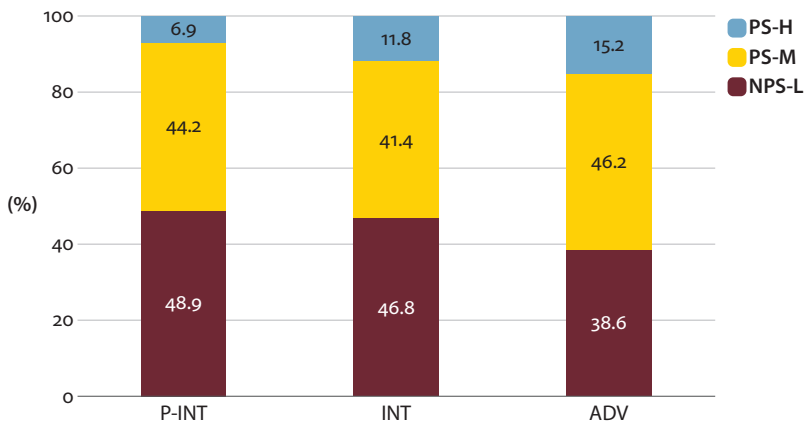
Figure 1. Depth of processing across proficiency levels

In particular, as shown in Figure 1, the percentage of low, non-problem-solving episodes decreases as proficiency increases, from almost half of the LREs produced by Pre-intermediate (48.9%) and Intermediate (46.8%) learners, to just 38.6% of the episodes generated by Advanced learners. Within the LREs that involved problem solving, those that presented a medium level of engagement were the most frequent for the three groups (between 44.2% and 46.2%), while high, problem-solving LREs were the least frequent. While they represented 6.9% of Pre-intermediate learners' LREs, the percentage increases for both Intermediate (11.8%) and Advanced (15.2%) participants. In terms of the number of episodes produced, Table 2 shows that the Intermediate participants engaged in twice more PS-H episodes than the Pre-intermediate group, while Advanced student-writers generated three times more episodes of this kind.

Research question 2

Our second research question asked whether the strategic orientation of our participants' problem-solving behavior varied as a function of L2 proficiency. Results indicate that orientation did not seem to affect the DoP of the episodes produced by Pre-intermediate and Advanced learners, while it appeared to have an effect on the DoP of Intermediate participants' LREs (see Table 3). Specifically, Pre-Intermediate participants' problem-solving episodes, irrespective of their DoP, were prompted mostly by problematic issues (compensatory orientation), while almost 75% of PS LREs observed in the Advanced participants' data were caused by their drive to improve their linguistic expression (upgrading orientation). In contrast, while most PS-M produced by intermediate learners were upgrading in nature, the majority of LREs that involved high levels of processing were caused by a compensatory goal.

Table 3. Strategic orientation of problem-solving LREs across proficiency levels

PS-M	P-INT			INT			ADV		
	N	M	SD	N	M	SD	N	M	SD
Upgrading	14	2.00	3.00	82	11.71	5.15	125	17.86	9.62
Compensatory	100	14.29	6.52	51	7.29	3.73	42	6.00	4.16
PS-H	N	M	SD	N	M	SD	N	M	SD
Upgrading	0	0.00	0.00	9	1.29	1.11	44	6.29	5.59
Compensatory	19	2.71	2.06	35	5.00	3.74	16	2.29	1.60

As shown in Figure 2, all the PS-H and the vast majority of PS-M LREs (87.7%) produced by Pre-intermediate learners had a compensatory orientation. The opposite was true for Advanced participants, who engaged in problem-solving behaviour mainly for upgrading purposes, irrespective of the depth of processing involved (PS-M = 74.9%; PS-H = 73.3%). In contrast, Intermediate participants' orientations for PS-M and PS-H LREs were quite different. While many PS-M LREs were upgrading in nature (61.7%), the majority of the episodes which displayed deep problem-solving mechanisms had a compensatory purpose (79.5%).
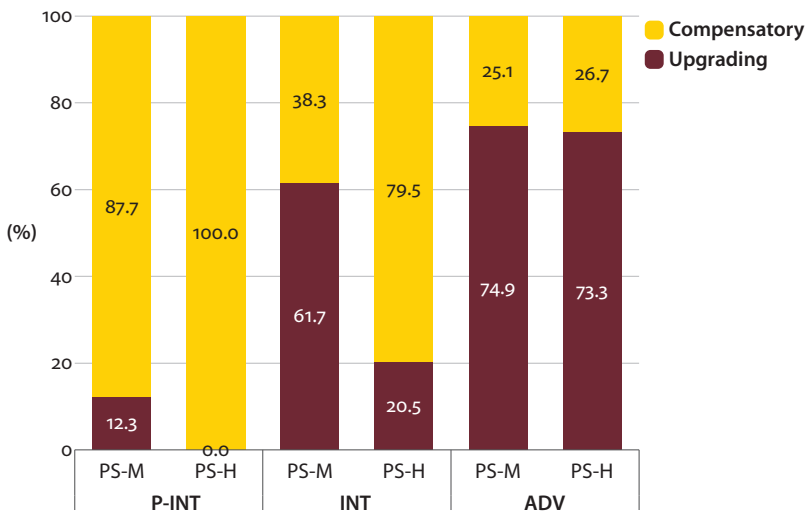
Figure 2. Strategic orientation of problem-solving LREs across proficiency levels

Discussion

The present study sought to contribute to the research agenda proposed by Cumming (1990) with its focus on the mediating effect of L2 proficiency (taken as availability of L2 knowledge) on learners' language reflection during individual writing. In particular, it is the first study to shed light on DoP and problem-solving orientation of linguistic processing, two constructs that were developed after Cumming's original proposal and that are purported to play a role in bringing about language learning through writing. Additionally, the study applies the theoretically-motivated coding system developed by López-Serrano et al. (2019) to the analysis of language-related episodes.

The theoretical and methodological decisions taken in the present study have allowed us to end up with a set of findings that we believe may be of relevance in understanding the role played by DoP in promoting language development through L2 writing. Theoretically, the study has been framed within Leow's ISLA model (Leow, 2015), which conceives of language learning as a process comprising (1) the detection and noticing of both content and linguistic data the learner is exposed to (input processing); (2) the establishment and/or strengthening of form-meaning connections through analyses and reanalyses of existing knowledge as well as for-mation and testing of hypotheses (intake processing); and (3) the use by the learner of the output produced as self-generated input or as an occasion to monitor what has just been produced (L2 knowledge processing).

While most research framed in the model has so far looked into the earlier stages of the L2 learning process, that is, input and intake processing, only one study has focused on the knowledge processing stage (Bergsleithner, 2019). Coincidentally, higher levels of processing tended to occur in Bergsleithner's and our own study when learners addressed their language concerns, while lower levels mostly appeared when they were able to access linguistic items that, although not fully automatized, were almost immediately available. On these occasions there was no need for learners to process their L2 knowledge with much effort and they managed to produce their text with accuracy.

However, it must also be noted that there are two relevant methodological differences between Bergsleithner's (2019) study and the present one that directly affect how the notion of DoP may be conceptualized in relation to writing: the type of tasks used and the operationalization of prior knowledge and L2 proficiency. In line with most studies based on Leow's (2015) ISLA model, Bergsleithner made use of a controlled task whose instructions asked participants to use a specific list of verbs in the past tense to create a narrative. Such a task hence prioritized the production of those verbs and tenses instead of focusing on what Cumming (1990) considered to be more conducive to language learning: The perception that the aim

of the task is to create a meaningful text that conveys information to the reader. In contrast, the participants in our study were asked to complete a meaning-oriented argumentative task, in which the grammatical and lexical items needed to formulate the intended message were not pre-established by the researcher but depended on the participants' own writing goals. Consequently, the more controlled nature and specific linguistic focus of Bergsleithner's task may have led learners to direct their attentional processes towards specific linguistic items and, as a result, may have limited the potential learning effects to just those elements. In contrast, the range of concerns addressed in our learners' attempts to generate their L2 texts resulted in the kind of multidimensional linguistic processing (noticing, reflection of language, etc.) that is claimed to promote learning in terms of consolidation and/or expansion of linguistic repertoires, analysis of explicit knowledge, or the gradual transformation of explicit knowledge into automatized knowledge (Manchón & Williams, 2016).

Another important difference between the two studies relates to the operationalization of prior knowledge and L2 general proficiency. The participants in Bergsleithner's study were drawn from the same level of proficiency given that the focus of the study was to address the role of prior knowledge on subsequent performance. In consonance with the type of task used, Bergsleithner conceptualized prior knowledge as recent or old knowledge of the English irregular past tense. In the present study, prior knowledge is embedded in a recent conceptualization of L2 proficiency (Hulstjin, 2011, p. 230) which involves two kinds of ability: (1) basic language cognition, which comprises explicit and implicit knowledge of lexical items and grammatical structures together with the automaticity with which these knowledge types can be processed; and (2) higher language cognition, which comprises lexically and grammatically less frequent and more complex utterances.

Bearing these assumptions in mind, higher levels of L2 proficiency should be understood in terms of the accuracy with which the learners at these levels process both basic utterances and less common linguistic elements (core components of language proficiency) with the use of metalinguistic knowledge, knowledge of oral and written discourse types and strategic competences (peripheral components of language proficiency). Our participants' L2 proficiency was assessed through the OPT, a test that allegedly measures language skills more than linguistic knowledge (Macaro, 2014) and, therefore, that assumes the application of the peripheral components of proficiency to the core ones if higher levels of proficiency are to be reached. That may have been the case with our advanced participants and, as a result, their deeper levels of processing were mainly geared towards the formulation of sophisticated vocabulary and grammatical structures. By the same token, as proficiency decreased, higher levels of processing were activated by learners to deal with more basic vocabulary and structures.

Concerning the effects of L2 proficiency on DoP, our results diverge from those reported in research on input processing. For instance, Calderón (2013) found that while performing a multiple-choice recognition task, the more advanced group engaged in less processing than their less proficient counterparts. In contrast, our data indicated that, in the production of an argumentative composition, i.e., a different task type, deeper processing was more frequent as proficiency increased. We may thus hypothesize that having more linguistic resources to draw on allowed our advanced participants to engage in deeper processing to find better ways to express their ideas in the L2, as will be discussed below in relation to our second research question. In the case of lower-level learners, whose L2 knowledge or access to it is more limited and who may suffer from cognitive overload, argumentative tasks may not be the most appropriate tasks to take full advantage of the language learning potential of writing. It would thus be relevant to test these hypotheses with studies that look into the combined effects of task-type and proficiency in promoting DoP and bringing about language learning through writing. On the other hand, in relation to the stages of the learning process proposed in Leow's ISLA model, our results suggest that producing output in the L2 (knowledge processing stage) may be more challenging and more cognitively effortful than processing incoming L2 data (input processing stage). However, more research is needed to test this hypothesis and to gain further knowledge on the potential role played by DoP at the different stages of the learning process.

Regarding our second research question, the data show that the notion of DoP as a cognitive mechanism likely to promote language learning in L2 writing would be more fully understood if it is seen in conjunction with the notion of "strategic orientation". Including this dimension in the analysis has allowed us to begin to understand what motivates learners at different proficiency levels to engage in language-related episodes that involve deep processing. In addition, as different orientations affect the strategies employed and the type of knowledge accessed by writers, they are likely to have an impact on the potential for language learning of the LREs.

Thus, our data show that pre-intermediate and intermediate learners engaged in deep output processing mostly for compensatory purposes, i.e., they generally engaged in search processes aimed to find ways to express their intended meanings in the L2. In doing so, they noticed holes and gaps in their L2 linguistic resources, formulated hypotheses through the activation of previous L1 and L2 knowledge, and evaluated the output produced as self-input through metalinguistic reflection, all of which are part and parcel of DoP (Leow, 2015). Nevertheless, they also found themselves limited by their prior knowledge and, therefore, their chances of consolidating partially acquired meaning-form relationships or creating new ones were reduced. In this sense, compensatory episodes would support the claim that while

deeper levels of processing may increase the likelihood of achieving higher levels of awareness as understanding, this is not always the case (Leow, 2015). Example (4) below is provided as an illustration of this phenomenon. INT2 engages in a PS-H compensatory search for an expression in the L2 that allows her to convey the idea, formulated in Spanish, that parents "set the example" or "are roles models" for their children. In her search, she retrieves alternatives in the L1 ("los conducen" (they lead them); "son un ejemplo a seguir" (set the example); "tienen mucha influencia" (exert a strong influence)). However, she is only able to access the L2 concepts "are an example" and "a mirror", which she writes down.

(4) PS-H compensatory
     Family is the center of their life until they go to school and it's obvious parents um (…) um ¿cómo podría decir que los (…)? a ver […] ¿cómo es que los conducen? (…) *family is the center until they go to school and it's obvious parents* **are** (…) (6) ay no sé cómo ponerlo (…) *parents are* (…) […] an example (6) en español sería un ejemplo a seguir pero es que en inglés (…) an example (…) a ver (6) and (…) no sé lo que poner voy a poner (…) mirror (3) no sé *family is the center of their life until they go to school and it's obvious parents are* (…) bueno lo que sea (…) y si no lo pongo en español y luego […] *it's obvious parents* erm (…) tienen mucha influencia pero (…) no es eso (…) *parents* um (…) bueno […] **an example and a mirror**   (INT2–E5)

The language learning potential of these episodes is to be found in the possibility they offer learners to go on, albeit imperfectly, with the writing task and potentially make them aware of new problematic areas in their L2 abilities. Perhaps more importantly, the learning potential can also be found in the traces that the cognitive effort involved in the activation of L2 knowledge (especially at higher levels of processing) may leave in learners' memory. If these traces remain or are activated by the time feedback is provided (see Leow, Chapter 5, this volume), learners may greatly benefit from either testing their hypotheses about the L2 or from receiving correct L2 forms in the written corrective feedback provided. This conjecture, which is consistent with the purported connectedness of input and output processes in SLA (Gass, 2010), should thus be tested in future studies on the processing of incoming feedback (see Chapter 17, this volume).

In turn, even though our advanced participants possessed greater control of the L2 than the two lower proficiency groups, this did not mean that they did not process their output deeply. Instead, they actively engaged in problem-solving processes that involved high depth of processing in order to improve (upgrading orientation) their texts. In these episodes they reprocessed their output with the intention of engaging in what Cumming (1990) called "search for improved phrasing": They repeatedly accessed their mental lexicon to retrieve forms that represented a better

match to their communicative intention. For instance, to avoid lexical repetition in their texts, our advanced writers activated searches for alternative expressions and synonyms. In other cases, such as when attempting to ensure the truth value of their statements, they looked for the wording that best captured their intended meaning, such as introducing expressions of modality instead of making simple affirmative statements, as can be seen in Example (5) with the use of "may", "maybe" and "probably":

> (5)   PS-H upgrading
>   nevertheless it is true that if you're not encouraged at home **nevertheless it is true that** the less **the less you are** encouraged (3) *it is true* and how do we know that it is true? no *it is true* no because there are cases in which it is not true [CROSSES OUT it is true that the less you are] (4) it's probably (4) *nevertheless* (3) the encouragement and support you receive at home **the encouragement and support** you may receive no **you receive at home** maybe **maybe influences** no motivates no influences no [CROSSES OUT influences] motivates you **motivates you more** […]   (ADV1-E44)

In this kind of episodes our participants explored the relationship among different terms or structures in their L2 or even their L1, therefore potentially allowing for the strengthening of connections between those expressions and perhaps opening up new, unsuspected avenues of form-function mappings along the way. In other words, while writing challenging texts, L2 writers might not only be consolidating their L2 knowledge but likely also establishing new, still underdeveloped form-meaning spaces for future development.

In short, our data show that the lower the level of proficiency, the more the participants in the study were involved in deeper processing for the purposes of tackling basic vocabulary and grammatical structures, although this cognitive effort toward searching for more sophisticated vocabulary and complex syntactic structures was reoriented as they moved up in the proficiency scale. This pattern indicates that the potential role DoP may play in promoting language learning via free writing tasks is not a black and white issue, but one that can be better understood if it is contemplated in connection with the strategic orientation (compensatory or upgrading) adopted by the writers when addressing their self-generated problems. We consider this to be a relevant additional piece to the puzzle that depicts the picture of the language learning potential of L2 writing. Our conjecture would be that at least an intermediate level of proficiency seems to be needed to benefit from writing complex meaning-oriented composition tasks in terms of language learning. In the case of intermediate learners, these tasks may play a double role: First, medium levels of processing are mostly activated to improve one's own texts, which, with frequent practice, may help learners consolidate the language used.

Second, their struggles to compensate for the gaps in their knowledge, which evidence deep processing levels, resemble those processes considered by Laufer and Hulstijn (2001) as presenting a high involvement load: Learners feel the *need* to find an expression in the L2, conduct intense *searches* within their interlanguage, and *evaluate* their available options. These searches and evaluations, conducted at high levels of processing, may present an optimal opportunity for noticing gaps and for acquiring new knowledge through their use of external resources or through directing their attention to incoming corrective feedback.

While the opportunities for language development for intermediate learners would mostly depend on their access to external sources, more advanced levels of proficiency may allow L2 writers to "evaluate forms of the L2 in relation to their intended meanings, search earnestly to find the best words to express ideas, and switch purposefully between languages to make principled decisions" (Cumming, Chapter 2, this volume, p. 32). Further research should thus be conducted to test these hypotheses through the longitudinal investigation of learners' engagement in repeated writing practice and through measures of their language development in relation to such tasks.

Conclusions, limitations, and future directions

Given the scarcity of research on individual composition writing in relation to its potential for language learning, the present study was intended as a contribution to the research agenda suggested by Cumming (1990) with a novel exploration of the mediating role L2 proficiency may have on learners' depth of processing and orientation of strategic behavior during the completion of an argumentative writing task. We believe the study has shed new light on the intricacies of the learning potential that may derive from L2 writing and has led to proposing a set of empirical questions worth addressing in future research agendas.

Despite this potential contribution, our study is limited given that it included a low number of participants, only one task, and one main data source. Furthermore, in order to obtain homogeneous groups in terms of L2 proficiency, our participants' ages ranged from 16 to 24. This meant that their educational and literacy experiences in English also differed across groups. In particular, the participants' writing ability was not controlled for and, as a result, the influence of this variable was assumed in an *ad hoc* fashion, as part of the participants' level of proficiency. Before more warranted statements about the role of DoP in the L2 writing process can be made, future research, whenever possible, ought to control for the independent contribution of L2 knowledge and writing ability.

Future studies must also include more tasks and genres, particularly since task demands appear to influence the kind of attention paid to language. As Cumming (1990) pointed out, more cognitively demanding tasks of an argumentative nature (such as the one used in the present study) tend to elicit higher instances of language reflection than less demanding tasks such as letters or brief data reports (see also Révész, Kourtali, & Mazgutova, 2017). Thus, it could be posited that the language learning potential of writing may be contingent to task type, although this is still a hypothesis that needs to be tested (but see Zalbidea, and Vasylets, Gilabert, & Manchón, this volume, in relation to the variable of task complexity). Another limitation of the study is its reliance on the participants' TA protocols as the main data source. Though verbalizations (either concurrent or retrospective) are still the only method available to capture certain writing behaviours (such as the strategic orientation adopted by writers) (Galbraith & Vedder, 2019), they should be combined with other modalities of data (i.e., eye tracking, key-stroke logging, digital screen capture, etc.) to obtain a more complete picture leading to a better interpretation of L2 writing processes (see Révész & Michel, 2019; Manchón & Leow, Chapter 14, and Stiefenhöfer & Michel, Chapter 11, this volume).

Despite these limitations, we consider our study a relevant contribution to current understandings of learners' reflection on language during individual composition writing tasks as, to our knowledge, it constitutes the first attempt to conceptualise DoP during written production of an open, complex writing task (an argumentative composition). Studies that capture L2 users' thinking processes are crucial to better understand how to guide them to make the most of their engagement with language problems during writing tasks. Further research on the role of DoP in language learning through writing should set up a research agenda with questions related to how, when, and why DoP may be associated to what kind of learning.

Acknowledgements

Funding

References

Allan, D. (1995). *Oxford Placement Test*. Oxford: OUP.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review* 89(4), 369–406. https://doi.org/10.1037/0033-295X.89.4.369

Bergsleithner, J. M. (2019). The role of prior knowledge in depth of processing during written production: A preliminary investigation. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 104–120). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-8

Byrnes, H., & Manchón, R. M. (2014). Task-based language learning: Insights from and for L2 writing. An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 1–23). Amsterdam: John Benjamins.

Calderón, A. M. (2013). The effects of L2 learner proficiency on depth of processing, levels of awareness, and intake. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honour in Richard Schmidt* (pp. 103–121). Honolulu, HI: University of Hawai'i.

Craik, F. I. M. (2002). Levels of processing: past, present, and… future? *Memory*, 10, 305–318. https://doi.org/10.1080/09658210244000135

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511. https://doi.org/10.1177/0741088390007004003

Galbraith, D., & Vedder, I. (2019). Methodological advances in investigating L2 writing processes: Challenges and perspectives. *Studies in Second Language Acquisition*, 41(2), 633–645. https://doi.org/10.1017/S0272263119000366

Gass, S. (2010). Relating L2 input and L2 output. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 194–219). London: Continuum.

Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second language assessment. *Language Assessment Quarterly*, 8(3), 229–249. https://doi.org/10.1080/15434303.2011.565844

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26. https://doi.org/10.1093/applin/22.1.1

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centred approach*. London: Routledge. https://doi.org/10.4324/9781315887074

López-Serrano, S., Roca de Larios, J., & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically-motivated and empirically-based coding system. *Studies in Second Language Acquisition*, 41(2), 503–527. https://doi.org/10.1017/S0272263119000275

Macaro, E. (2014). Reframing task performance: The relationship between tasks, strategic behaviour, and linguistic knowledge in writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 53–77). Amsterdam: John Benjamins.

Manchón, R. M., & Roca de Larios, J. (2007). Writing-to-learn in instructed language learning contexts. In E. Alcón-Soler & M. P. Safont-Jordá (Eds.), *Intercultural language use and language learning* (pp. 101–121). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-5639-0_6

Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing. Implications for theory. In R. M. Manchón (Ed.), *Foreign language writing: Learning, teaching and research* (pp. 102–124). Clevedon, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859-008

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón, & P. K. Matsuda (Eds.) *Handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97. https://doi.org/10.1037/h0043158

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, 37(3), 439–468. https://doi.org/10.1111/j.1467-1770.1987.tb00579.x

Révész, A., Kourtali, N., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, 67, 208–241. https://doi.org/10.1111/lang.12205

Révész, A., & Michel, M. (2019). Introduction. *Studies in Second Language Acquisition*, 41(2), 491–501. https://doi.org/10.1017/S0272263119000329

Roca de Larios, J., Manchón, R. M., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behavior in the allocation of time to writing processes. *Journal of Second Language Writing*, 17(1), 30–47. https://doi.org/10.1016/j.jslw.2007.08.005

Shing, Y. L., & Brod, G. (2016). Effects of prior knowledge on memory: Implications for education. *Mind, Brain, and Education*, 10(3), 153–161. https://doi.org/10.1111/mbe.12110

Storch, N. (2016). Collaborative writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 387–406). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-021

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–153). Rowley, MA: Newbury House.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step toward second language learning. *Applied Linguistics*, 16, 371–391. https://doi.org/10.1093/applin/16.3.371

Investigating the relationship between peer interaction and writing processes in computer-supported collaborative L2 writing

A mixed-methods study

Laura Stiefenhöfer and Marije Michel

Lancaster University / Groningen University

Twenty-first century technology has created new digital contexts (e.g., shared online writing platforms) that influence the potential of collaborative writing for second language (L2) learning. Making a methodological contribution to studying L2 computer-supported collaborative writing (CSCW), the present chapter reports on an exploratory yet innovative investigation triangulating data from text mining, interaction analyses, eye-tracking, and stimulated recall. Eight international students in the UK used Google Docs for paired collaborative writing tasks. Results of quantitative and qualitative analyses reveal the many ways in which peers interact with each other and the nature of the emerging text during CSCW. Findings illustrate the complexities of CSCW and indicate how triangulating different methods facilitates the study of the affordances of CSCW and its potential contribution to L2 learning.

Introduction

The past three decades have seen a body of research that provides ample evidence for the ways in which writing in a second language (L2) leads to language learning (Cumming, 1990; Manchón, 2011; Manchón & Roca de Larios, 2007; Roca de Larios, Nicolás-Conesa, & Coyle, 2016). This previous work identified two main features of writing that are potentially conducive to L2 learning: (1) the slower pace of written production, which allows learners to exploit different kinds of knowledge; and (2) the permanence of the written output, which provides opportunities to compare, revise, and notice gaps in the emerging text (Manchón & Williams, 2016). To a lesser extent, L2 collaborative writing and its role for second language

acquisition (SLA) has also received some scholarly attention (Storch, 2009). For collaborative L2 writing, peer interaction between writing partners and the opportunity of pooling linguistic resources of all contributors are central features explaining why, particularly, these activities might benefit language learning (Storch, 2016).

Due to technological advancements of the current century, new digital contexts have emerged (e.g., wikis and shared online writing platforms) that have transformed the socio-cognitive environment of collaborative writing, be it in L1 or L2 (Yim & Warschauer, 2017). The affordances of digitally-mediated collaborative writing have sparked new work examining, for example, L2 learners' interactions in online contexts (Rouhshad & Storch, 2016). The study reported in the present chapter aims to make a methodological contribution to this line of research by studying peer interaction during L2 computer-supported collaborative writing (CSCW) using Google Docs. Notwithstanding its small-scale and exploratory nature, by using an innovative combination of text mining tools, interaction analyses, eye-tracking, and stimulated recall interviews, our study sheds light on the affordances of CW in digital environments as a site for L2 learning.

In the following, we will first review work addressing the value of collaborative writing for L2 learning with a specific focus on interactional patterns during digitally-mediated writing activities. We will then highlight some methodological innovations and challenges of studying CSCW.

Collaborative writing in an L2 in a digital age

The language learning potential of collaborative writing

Earlier work on L2 writing showed that text composition processes in the L2 are similar to those in the L1, but more time consuming and challenging due to lower accessibility and incompleteness of L2 knowledge and the increased efforts needed for formulation (Révész, Michel, & Lee, 2019; Roca de Larios, Manchón, & Murphy, 2006). In this light, collaborative writing (CW), that is, two or more writers co-authoring a text (Storch, 2011), might present a unique means for L2 writers to cope with the high cognitive demands of L2 writing as it enables them to draw on the joint knowledge and linguistic resources of all co-authors (Storch, 2013). While most work on peer interaction has focused on oral communication (Mackey, 2007), L2 writing research indicates that the written modality may create distinct language learning opportunities during CW. For example, providing and receiving feedback and noticing errors (Manchón & Williams, 2016) might be supported by the fact that written output is more permanent and salient than oral output, allowing learners to address linguistic concerns with fewer online processing constraints

(Harklau, 2002). Moreover, Bazerman (2016) argues that CW encourages learners to verbalise many of the implicit steps of the writing process when interacting with their partner. Consequently, planning, revising, as well as providing feedback to their partner's writing potentially improve L2 writers' composing process. Similarly, receiving immediate feedback by a peer as partners work on a joint text is likely to be more effective than the (delayed) feedback on individual writing provided by a teacher or peer (Polio, 2012).

CW activities may be used in L2 pedagogy to foster specific professional or academic writing skills (Storch, 2017) or with the general aim to improve L2 competence, a concept Manchón (2011) coined as "writing-to-learn language" (WLL). Hence, L2 CW has some characteristics which previous SLA research has linked to language learning, in this case WLL. First, CW activities provide ample opportunities for and may even push learners to produce output, which is in line with Swain's Output Hypothesis (Swain, 2000). Second, CW encourages learners to engage in meaningful interaction (Izumi, 2002), with all the favourable language learning features highlighted by the interactionist approach, such as receiving and providing feedback, modifying input and output, or negotiation for meaning and form (Long, 1996; Mackey, 2007). Both of these characteristics are thought to focus learners' attention on form, which, presumably, increases noticing (Schmidt, 2001). Third, Storch (2016) argues that CW presents learners with opportunities to engage in languaging, a socio-cultural concept that refers to the use of language to make meaning, in order to make sense of complex information and tasks (Swain, 2000, 2006). Languaging in CW is externalised, thus providing L2 writers with opportunities to pool their knowledge together while making use of collective scaffolding (Storch, 2016). Donato (1994) sees collective scaffolding in the L2 (i.e., the assistance given to each other by learners of both different and similar proficiency) as a means for learners to extend their "current skills and knowledge to higher levels of competence" (p. 40). Therefore, CW potentially enables learners to perform at a level beyond their individual L2 competence and capabilities. In sum, CW is thought to support L2 learning through both the potential of writing itself to learn language, and the benefits of the interaction that is criterial to CW (Storch, 2013).

Patterns of interaction in collaborative writing

Storch's (2001, 2002, 2009) seminal research has taken a close look at interactional processes during CW. Investigating the types of relationships that learners build when writing collaboratively, she developed an analytic framework to classify interactional patterns drawing on Damon and Phelps' (1989) indices of equality and mutuality. Damon and Phelps (1989) describe equal interactions as those where "both parties in an engagement take directions from one another rather than one

party submitting to a unilateral flow of direction from the other […]" (p. 10). Mutuality reflects the extent of the engagement with the partner's contribution. Highly mutual interactions include a large number of instances of reciprocal feedback and co-construction (Storch, 2013). Storch's model distinguishes four different patterns: (1) Collaborative interactions emerge when both participants show equal engagement with a task and also demonstrate similar levels of involvement with their partner's contributions; (2) Cooperative interactions (or dominant/dominant) are characterised by similar amount of contribution to a task (i.e., high equality), but lacking engagement with the partner's input (i.e., low mutuality, (Storch, 2002); (3) The expert/novice-pattern (low equality, high mutuality) is marked by a (perceived) asymmetry regarding competence, which results in the expert assuming greater responsibility for the task while still encouraging the novice to contribute; (4) In the dominant/passive pattern, the dominant participant, who is not necessarily more competent than their partner, "seems to appropriate the task" (Storch, 2002, p. 129), without engaging with their partner.

Earlier empirical work suggests that the collaborative and expert/novice pattern are the most beneficial for language learning because they elicit more languaging and more instances of collective scaffolding (Kim & McDonough, 2008; Storch & Aldosari, 2013). Pairs adopting these patterns also show more evidence of knowledge transfer and language uptake (Storch, 2002; Watanabe & Swain, 2007). Research has also explored a number of variables potentially influencing interaction, such as L2 proficiency (Kim & McDonough, 2008), group size (Fernández Dobao & Blum, 2013), and mode of communication (Rouhshad & Storch, 2016; Tan, Wigglesworth, & Storch, 2010).

To sum up, the available work shows that Storch's model provides a feasible framework to classify patterns of interaction during CW activities, even though some contexts might require adaptations. An important development that is likely to influence writing partners' interaction is that nowadays most CW takes place in web-based environments, such as Google Docs (Godwin-Jones, 2018). Accordingly, the next section reviews the specific affordances of CSCW and how these relate to interactional patterns.

L2 writing in CSCW

Web applications such as wikis, social networking sites, and Google Docs have become essential platforms for CW as they allow people to engage in joint projects and interaction regardless of their time and location (Yim & Warschauer, 2017). Research into CSCW is growing fast (as reviewed in Li, 2018) and an increasing number of software tools for the analysis of CSCW offers new insights into CW processes (Yim & Warschauer, 2017).

Online collaborative writing platforms have some characteristics that make them specifically relevant for L2 pedagogy. One of the main affordances of CSCW is its transparency regarding the writing and drafting process of each writer involved. In Google Docs, co-authors cannot only access the document history to compare different versions of a text, but even observe the writing in real time. It also includes a written chat tool, which co-authors can use to interact with each other. The interaction as well as the emerging text is visible on the screen and remains accessible throughout the whole composing process.

Previous research on written computer chat has argued that the salience of digital writing is particularly supportive of L2 learning (Sauro, 2009; Smith, 2005). The permanence of output (written composition and text chat conversation) may serve as a source for vocabulary and grammar, available for cognitive comparison and subsequent writing (O'Rourke, 2008). Additionally, also an observing partner might develop their L2 as the digital environment promotes attention to form (O'Rourke, 2012; Smith, 2005). It is likely that this also holds for CSCW.

Only a few empirical studies have investigated interactional patterns during CSCW. For instance, Tan et al. (2010) found that both modes (face-to-face [FTF] vs. computer-supported chat) elicited a similar amount of collaboration. Yet, only the CSCW mode triggered cooperative interaction, with little evidence of collaborative scaffolding, while the FTF interaction prompted a lower degree of equality with more expert/novice and dominant/passive patterns. Similarly, Rouhshad and Storch (2016) compared computer vs. FTF chat during CW and showed that most pairs displayed cooperative and dominant/passive patterns in text chat, while partners in the FTF condition often adopted a collaborative approach. A follow-up study revealed large individual differences in the ways in which participants engaged in negotiation for meaning, probably due to the extra processing time available in text chat (Rouhshad, Wigglesworth, & Storch, 2016). While Storch (2016) suggests that text chat promotes symmetrical relationships and more equal participation, she reminds us that equality does not necessarily result in engagement with the partner's contributions, as writers sometimes seem to only focus on their own part of the task. Cho (2017) gives further support to this interpretation, as her participants adopted more collaborative patterns when using voice chat (instead of text chat) during CSCW. In addition, other research on interaction in CSCW has highlighted the role of the writing environment on the emerging patterns (Li & Zhu, 2013) and has expanded Storch's categories with additional labels (Abrams, 2016; Cho, 2017; Li & Kim, 2016).

Other CSCW work has investigated the quantity and quality of revisions writers make to their partners' text as a form of collective scaffolding (Storch, 2002, 2005). Findings indicate large individual differences in the amount of edits amongst learners (Kessler, Bikowski, & Boggs, 2012) and different prevalence for different types

of revisions (i.e. meaning making or formal revisions) depending on the writing environment (e.g. wikis and Google Docs) (Arnold, Ducate, & Kost, 2012; Kessler et al., 2012; Li & Zhu, 2013).

In sum, earlier work suggests that the environment for CSCW and communication influence the interactional patterns that might emerge during collaborative L2 writing activities and that writers differ in how much and what they edit in their partners' texts. Yet, more work is needed to further our understanding of how these features of CSCW lead to language learning.

Crucially, from a research methodology perspective, all the studies reviewed so far used methods that provide insights into what L2 writers actively contributed to a collaborative task in terms of writing, editing, or interacting with their partner. However, visual attention to different parts of the screen can serve as an additional indicator of the participants' engagement with the digital environment (Stickler & Shi, 2015). As such, measuring gaze behaviour during CSCW can provide valuable information about interactional patterns. Moreover, it can serve as a window into observational learning (i.e. learning from models) (Braaksma, Rijlaarsdam, van den Bergh, & van Hout-Wolters, 2004) during CSCW. The next section reviews earlier work using eye-tracking methodology to study online L2 learning, highlighting insights and challenges that might be associated with measuring gaze behaviour during CSCW.

Eye-tracking methodology when studying collaborative writing: Insights and challenges

Eye-tracking methodology, the colloquial name for eye-gaze measurement, builds on the assumption that visual attention to information on a screen in terms of focus, order, and duration of eye fixations gives insights into what information is being processed by a viewer's mind (Reichle, 2006). In short, eye-gaze data reveal which areas on the screen an individual has looked at and processed while carrying out a task (Poole & Ball, 2006). In recent years, eye-tracking has become popular as a method in applied linguistics research (Conklin & Pellicer-Sánchez, 2016) including work looking at reading during writing (Révész et al., 2019). However, Michel and Smith (2017) note that there is a certain hesitation to use eye-tracking in the field of digitally-mediated language learning. One possible explanation for this, as pointed out by pioneers in the field working on text chat data (Michel & O'Rourke, 2019; O'Rourke, 2008, 2012; Smith, 2009), might be the highly dynamic character of online environments caused by, for example, scrolling up and down. Even though eye-tracking can provide rich information about the viewing behaviour in online environments, thereby overcoming disadvantages of other methodologies (e.g.,

reactivity of think-aloud protocols), the dynamic nature of digital environments poses specific methodological challenges that complicate data coding and analysis (Michel & Smith, 2017).

Some earlier eye-tracking work on text chat provides evidence that learners' noticing of recasts (measured by duration of eye-fixations) is positively related to pre-post-test gains on a grammar test (Smith & Renaud, 2013). Data by Michel and Smith (2018) suggest that the number of eye-fixations predicts lexical alignment, that is, whether there is uptake of words from a peer, which can be a sign of learning. Focusing on interaction, O'Rourke (2008, 2012) showed how eye gaze replay can reveal in which order learners attended to the incoming messages by a chat partner, providing a more valid perspective on how a chat conversation evolved.

This short review shows that eye-tracking methodology has the potential to provide additional insights into learners' interaction with peers and with the digital environment during CSCW. In particular, eye-gaze data can serve as indicator of less salient interaction, such as reading messages of the partner, monitoring, and observational learning via reading.

To the best of our knowledge, to date no research has used eye-tracking in CSCW. In line with earlier work into L2 individual writing (see studies in Révész & Michel, 2019), and in an attempt to provide further insights into the learning affordances of CW, the present study aims to showcase how combining different methodologies with eye-tracking allows us to get a fuller picture of the interactional processes L2 learners engage in during CSCW.

The present study

Research question

Our study pursued the following research question:

> What patterns of interaction emerge during computer-supported collaborative writing using Google Docs?

We draw on Storch's (2009) model of equality and mutuality as we investigate this question by looking at:

1. Information on writing and editing behaviour provided by the Google Docs text mining extension DocuViz;
2. Amount of contribution to and language functions of written chat messages;
3. Visual attention to different areas on the screen using eye-tracking methodology;

Triangulating these data with stimulated recall interviews, we aim to relate the interactional patterns to individuals' perceptions of the collaboration. Through this procedure, we hope to gain further insights into how CSCW might support language learning.

Method

Research context and participants

Participants were recruited from an English for Academic Purposes (EAP) pre-sessional course at a British university.

For the current study, eight students with an average age of 25.5 years (SD = 5.9) and a mean of 15.3 years (SD = 5.6) of previous English studies signed up on a voluntary base. The majority had spent a month living in an English-speaking country, and they had an overall IELTS score of 6.0 to 7.0 (6.0 to 6.5 for Writing) at the time of data collection. Six of them had signed up as pairs, the other two met the first time for the study.

As reported on a pre-task questionnaire targeting experience with digital tools, only one of the participants had employed Google Docs prior to data collection but, except for one, they all said to use online chat and messenger services for one to four hours a day. The other student reported less than one hour of online chat conversations per day.

Table 1. Description of research participants

Pair	Participant	Age	Degree	Gender	L1	Years of studying English	Months in English speaking country	IELTS score overall	IELTS score writing
1	Rui	24	Postgraduate	female	Mandarin	16	1	7	6.0
	Xiran	23	Postgraduate	male	Mandarin	10	1	6.5	6.5
2	Ning	25	Postgraduate	prefer not to say	Mandarin/ Cantonese	15	1	6.5	6.0
	Yisi	26	Postgraduate	female	Mandarin	12	1	6.5	6
3	Jingmin	19	Undergraduate	female	Mandarin	12	1	6	6
	Yanwei	22	Postgraduate	female	Mandarin	10	1	7	6
4	Arjun	26	Postgraduate	male	Bengali	22	1	6.5	6
	Kazuo	39	Postgraduate	male	Japanese	25	18	7	6.5

Note. All participant names are pseudonyms

Task

The task constituted a paired collaborative, decision-making writing task which they completed in Google Docs. Students were asked to write an email to a student asking for a recommendation on what country to choose for his studies. They were provided with information on economic aspects as well as living conditions in different countries to support their advice. This decision-making task thus included an element of data commentary (i.e. commenting on non-verbal material such as graphs and charts) which is an important text format in academic writing (Swales & Feak, 2012). This task could be completed in the 30 minutes that were available (Wigglesworth & Storch, 2009).

Data collection tools and procedure

Before data collection, the participants had one training session to familiarise themselves with the different tools. Pairs then signed up for their joint session that took 1.5 to 2 hours. During the experimental sessions, individuals were working in two separate rooms, each participant seated at their own computer that was set-up for the study. After providing informed consent, they filled in a questionnaire asking about prior experience with digitally-mediated language learning and demographic characteristics.

Participants were then provided with 10 minutes of individual time for task preparation and note taking in the online application Google Keep, using the main task instructions and prompts presented as an on-screen word document. Next, pairs had 30 minutes to write a joint text using Google Docs. All task elements (instructions and data) were presented in the same Google Docs document the participants were asked to write in. To allow for eye-gaze measurement, the document was pre-set to the monospaced *Inconsolata* font, size 18. For interaction during task completion, students could use the written chat modality in Skype, which was available to them on the same computer screen (cf. Figure 1 for the screen set-up).

Throughout the main task performance, both students' eye-gazes were tracked using the Tobii system (Tobii Studio 3.0.9). Due to available equipment restrictions, one of the participant's eyes were tracked with a Tobii TX300 integrated system (sampling rate = 300Hz; 23″ TFT screen), the other one's on a Tobii X2-60 (sampling rate = 60Hz; 17″ screen Dell laptop). Participants' eyes were calibrated just before starting the main task (after the questionnaire).

After the main task, students took turns in participating in individual stimulated recall sessions prompted by a replay of the screen recording of their task

performance overlain with their eye-movements. They received a short explanation on how to interpret the eye-gaze recordings. Then they were asked to report what they had been thinking during the task prompted by the gaze replay. If necessary, they could pause the recording to elaborate on their thoughts. In a second round, the researcher replayed and paused the video to ask about participants' recollections of specific instances that seemed to provide information about collaboration (e.g., a long eye-fixation on the chat contribution of the partner). Participants reported in English but were allowed to switch to their L1 if they encountered difficulties when expressing an idea. Stimulated recall interviews took about 30–40 minutes.

Data coding and analysis: Equality and mutuality

Data of the different sources were analysed in light of the framework on dynamics of peer interaction, drawing on concepts of "equality" and "mutuality" (Damon & Phelps, 1989), as used in Storch's model of dyadic interaction (Storch, 2002, 2009). The notion of equality refers not only to equal contribution to a task, but also includes the nature of decision-making processes and the degree of task control (Storch, 2013). Mutuality was operationalised as the participants' engagement with each other's contribution. Both the direct involvement with the partner's text through editing as well as the extent to which participants responded to each other in the chat interaction was analysed within this framework (Li & Kim, 2016).

Text generation and editing of (peer) text

The Google Chrome extension DocuViz (Wang, 2016) provided data on writing and editing behaviours. Specifically, we looked at:

a. students' contribution to text generation (i.e., words written);
b. equality of contribution (i.e., the proportion of words written for the final text; Yim, Wang, Olson, Vu, & Warschauer, 2017); and
c. engagement with the partner's text (i.e., the number of edits to text written by the partner, and total number of edits in text).

Chat interactions

Contributions to the written Skype chat interactions (i.e., the total number of words and turns per participant) were counted as a measure of equality. Using Atlas.ti8, a measure for engagement was established by coding for language functions in the chat interactions distinguishing initiations (questioning; requesting; stating;

suggesting) from responses (agreeing; disagreeing elaborating; requesting; suggesting) following (Li & Kim, 2016). A second rater also coded the chat interactions, resulting in one case of disagreement, which was resolved through discussion.

Visual attention as measured by eye gazes

For visual attention, Tobii Studio software (version 3.0.9) was used to mark the different elements on the screen as a separate area of interest (AoI, cf. Figure 1): graphical information on the task sheet; notes in Google Keep; writing pane of Google Docs; chat log – reading and writing – in Skype. When an AoI moved to a new location (e.g., through scrolling), it was deactivated and reactivated at the new location. Subsequently, information on the total fixation duration (in milliseconds) for each AoI was extracted as a measure of engagement with that particular part of the screen.
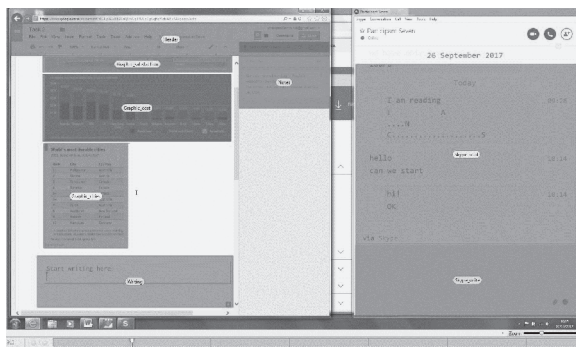
Figure 1. Definition of AoIs quantitative results

In the following, we will provide overviews of our quantitative results before we interpret and discuss our findings as we triangulate the information from the different data sources.

Contributing and editing during text generation

Table 2 shows the numbers for text generation and (peer) editing in the final text provided by DocuViz (Wang, 2016). There is substantial variation across participants regarding the number of words they contributed to the final text. Pairs also differed on the calculated measure for equality of participation (Yim et al., 2017).

For peer editing, Table 2 reveals substantial variation from 11 edits by Rui to 872 characters by Yisi, indicating different degrees of engagement.

Table 2. Equality and engagement in texts and chats (cf. DocuViz analysis)

		Texts				Chats		
		Number of words	*Equality of participation*	*Edit of other*	*Total edit*	*Turns*	*Words*	*Words per pair*
Pair 1	Rui	82	96.4	11	939	16	72	128
	Xiran	142		85	2449	15	56	
Pair 2	Ning	138	99.2	308	1572	7	46	159
	Yisi	179		872	1538	10	113	
Pair 3	Jingmin	68	85.1	406	1642	8	27	51
	Yanwei	231		135	5055	8	24	
Pair 4	Arjun	331	76.6	95	2584	15	154	237
	Kazuo	62		106	289	14	83	

Note. Equality of participation = (1 – variance of proportions of the total contribution) x 100. Higher scores indicate higher equality of written participation (Yim et al., 2017); Edit of other: Amount of editing each student did on a peer's text, measured by number of characters inserted or deleted within content contributed by the partner; Total edit: total number of inserted and deleted characters.

Contributing to and language functions within text chat conversations

In Table 2, figures for the chat interaction demonstrate that overall, chat conversations were fairly short (maximum of 31 turns) and that pairs varied regarding the quantity of contributions (from 51 words for Pair 3 to 237 words for Pair 4). Partners in all pairs produced a similar amount of turns, suggesting equality, but they differed considerably regarding the length of their contributions.

Table 3 summarises the language functions (cf., Li & Kim, 2016) that were used in the chat conversations. Most pairs showed balanced numbers of initiation with the exception of Pair 2, where Yisi initiated most interactions. Response patterns suggest higher values for Ning and Arjun compared to their partner, but given that chat interactions were short, figures are generally low here and must be taken with care.

Table 3. Language functions performed by participants

Language function	Participants							
	Pair 1		Pair 2		Pair 3		Pair 4	
Initiation	*Rui*	*Xiran*	*Ning*	*Yisi*	*Jingmin*	*Yanwei*	*Arjun*	*Kazuo*
Questioning	1	0	1	3	1	1	0	2
Requesting	1	3	0	0	0	0	2	0
Stating	2	1	0	2	1	1	2	1
Suggesting	1	1	0	1	0	1	0	1
Total	**5**	**5**	**1**	**6**	**2**	**3**	**4**	**4**

Table 3. (*continued*)

Language function	Participants							
	Pair 1		Pair 2		Pair 3		Pair 4	
Response	*Rui*	*Xiran*	*Ning*	*Yisi*	*Jingmin*	*Yanwei*	*Arjun*	*Kazuo*
Agreeing	0	2	0	0	1	0	0	2
Disagreeing	2	0	1	0	0	1	1	1
Elaborating	0	0	0	1	0	0	4	0
Requesting	0	0	1	0	0	0	2	1
Suggesting	0	0	3	1	0	2	0	0
Total	**2**	**2**	**4**	**2**	**1**	**3**	**7**	**4**

Eye-gaze data

Total fixation durations per Area of Interest (writing pane; graphical information of the task; chat input by the partner; own chat writing area; other) were added up for each participant and respective percentages of visual attention for each area are presented in Figure 2. Accordingly, participants differed considerably on the proportions of time they spent on different areas. Yet, they all fixated the longest on the writing pane (50% to almost 80%), followed by variable attention to the graphic sources (ranging from 2% to 23%), their partner's chat input (almost 4% to almost 20%), and their own chat writing box (just below 3% to more than 16%). Within each pair, one participant (Rui, Ning, Yanwei, Kazuo) spent substantially more time on the graphics than the other (Xiran, Yisi, Jingmin, Arjun). Similarly, one participant (i.e., Xiran, Yisi, Jingmin, Kazuo) gave the chat input more attention than the other.
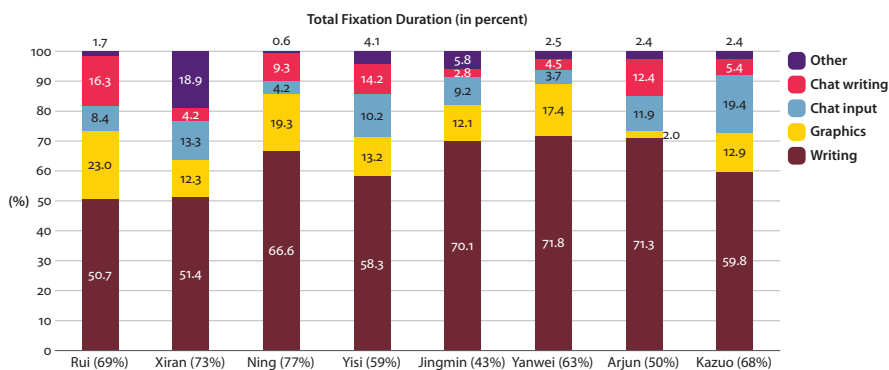
Figure 2. Total fixation duration in different AoIs (% gaze samples of each participant)

Data triangulation and discussion

As a final interpretational step, we triangulated the findings from the different data sources and complemented them by the stimulated recall comments in order to classify the interaction patterns of our participants drawing on Storch's (2009) model (cf. Figure 3). In the following, we will present a rationale for the classification of each pair as we discuss the theoretical and methodological implications of our study.
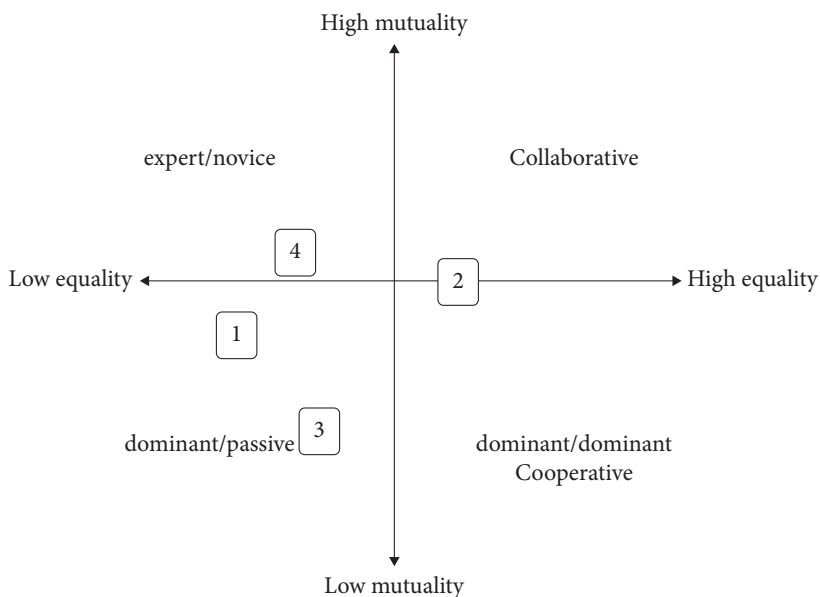
Figure 3. Identified interaction patterns

Classification of interaction patterns based on data triangulation –
The value of self-reports (Pair 1 – dominant/passive)

In the first pair, Xiran contributed and edited more of the final text, editing almost eight times more text than his partner Rui. Even though both partners interacted with almost equal number of turns during the chat interaction, Xiran's contributions were mostly initiations. He also dedicated more than 13% of this visual attention to the chat messages of his partner.

The stimulated recall comments reveal a large discrepancy between both partners' perception of the task. Xiran expressed a strong sense of single ownership ("my essay"), motivated by the fact that his partner admitted at the beginning that

"she had no clue about the assignment". He, on the other hand, was quite confident with his understanding of the task. In contrast, Rui, commented that she deliberately chose the more passive approach to avoid gratuitous work. She was waiting for Xiran to finish the main part and only started with the conclusion after his explicit request to do so. Accordingly, Pair 1 was classified as *dominant (Xiran) / passive (Rui)*.

To date, most research on interaction patterns has based the classification on participants' contributions to the final text and their engagement in talk during CW (e.g. Storch, 2013). Our findings show that complementing these data with the analysis of the chat interaction in terms of equality and editing as a measure of engagement seems to facilitate more valid classifications as it allows us to consider not only interpersonal interaction, but also interaction via the text itself. Yet, more and different data sources in our study also increased the complexity of the classification task: We needed to decide how to factor in each type of information, and how to deal with possible contradictions, for example, discrepancies between text contribution and chat content. Our findings suggest that the self-perceived role (as commented on during the stimulated recall interviews) constituted a crucial source of information that helped us in disambiguating contradicting data. As detailed below, the stimulated recall comments also provided valuable insights into perceived instances of language learning.

Classification of a highly dynamic process of interaction
during CSCW – The value of eye-gaze information
(Pair 2 – alternating between collaboration and cooperation)

For Pair 2, the computed measure of equality of participation is very close to 100, which indicates high equality. Ying and Nisi divided the task at the beginning. Even though Yisi was slightly more active, throughout task performance both engaged in discussions of alternative views, and both performed several edits to their partner's text. Drawing on the chat interactions, Yisi held the initiating role while Ning responded, but both participants were actively engaged in the task and the chat interactions. The eye-gaze data reveal that both gave more attention to writing their own messages than to reading the input of their partner. In particular, Ning dedicated more than 80% of her attention to the writing pane and graphical task information, suggesting that she was focusing on (her own) task. Based on these findings, we might place this pair either within the *collaborative* or the *cooperative* quadrant of Storch's (2009) model.

Again, the stimulated recall comments provided some crucial information. Both students mentioned how they came to a clear division of labour regarding

who was going to write what, suggesting cooperation (Storch, 2009). However, while they gave each other plenty of autonomy with their respective sections, they repeatedly checked their partner's writing to ensure text coherence, as exemplified in the following comment by Yisi: "I think I should check it [Ning's part] because I was writing the main body of this task and I think I should make sure the conclusion [written by Ning] and the main body is not conflicting with each other."

Furthermore, as demonstrated in Excerpt 1, this pair used the chat tool to resolve discrepancies between their respective views to come to a joint version of the text. Accordingly, interaction of Pair 2 could also be labelled as *collaboration*.

Excerpt 1. Chat interaction Pair 2

Time stamp	Part.	Text
[17:02]	Yisi:	the student are supported by family or himself
[17:03]	Ning:	not mentioned
[17:03]	Yisi:	i am thinking about this sentence
[17:03]	Yisi:	it might be better for you to discuss it together with your family
[17:05]	Ning:	what is the problem? it might have some grammar mistake?
[17:07]	Yisi:	no grammar mistake, just cultural differences. if the student is supported by his family, then he should discuss with them. But if he is self financed, then he don't need to discuss with others because he's 18 years old. He can have his own decision.
[17:07]	Ning:	but this is our advice we dont ask him to do
[17:08]	Ning:	it might be better for you to discuss it together with your family or friends. this better?

Halfway through the interaction presented in Excerpt 1, Ying presents a concern ("it might be better…"). The time stamps reveal that almost two minutes pass before Ning reacts to her. By taking a detailed look at the gaze-replay during this specific episode of the writing process, we learn that Ning reads Yisi's chat message right after it was sent and fixates for a long time on this message before she asks for further clarification. While waiting for Yisi's reply, Ning re-reads a question that Yisi had asked a minute earlier ("the student are supported by family or himself") (cf. Figure 4) and, eventually, goes back to the writing pane to edit a word in a different sentence than the one Yisi was asking about. In other words, she does not engage with the concern raised by Yisi in the first place. Consequently, this episode might be better classified as *cooperative*.

Based on the different sources, we concluded that Ning and Yisi formed a *collaborative-cooperative* pair.
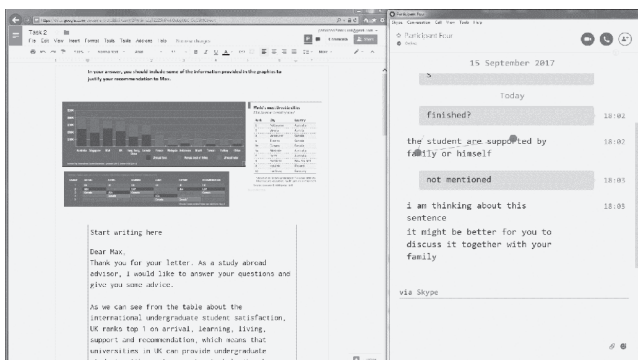
Figure 4. Eye-gaze replay

The challenging classification of Pair 2 illustrates the difficulty of assigning the static labels proposed by Storch (2009) to the highly dynamic patterns of interaction that emerge and change throughout a CSCW task (Eddy-U, 2015). In our case, we could work with Storch's labels by combining them. Other researchers have described different patterns that led to elaborations of the original model. For example, Abrams (2016) reports on a pair that was *passive/passive* and Cho (2017) added a *facilitator/participant* category. Not least, our findings reiterate the call by Young and Tedick (2016) who propose close examinations of single episodes, in order to identify dynamics of interaction at a micro level. It seems that the triangulation of several data sources (as done in the current study) is particularly valuable for a more detailed analysis, potentially, allowing for the identification of those patterns and dynamics that support language learning.

Affordances and limitations of communication during CSCW (Pair 3 – dominant/passive)

In Pair 3 we also see a clear division of labour: Yanwei contributed more text while Jingmin did more of the editing, including frequent edits to her partner's text, which resulted in a fairly low equality of participation. The short chat conversation demonstrates that these roles emerged naturally without the partners explicitly discussing it. In fact, the few instances where one partner initiated some sort of decision-making, the other did respond with a very short answer or not at all. Also the eye-gaze data show that both partners gave most attention to the writing pane and the graphics of the task, adding up to more than 80% (Jingmin) and almost 90% (Yanwei), leaving not much space for interaction. This lack of engagement with the partner paired with low equality suggests a *dominant* (Yanwei)/*passive* (Jingmin) pattern.

The stimulated recall comments support this classification. From the beginning onwards, Yanwei adopted the role of leading writer with her partner serving as spellchecker or editor. She wanted to work independently, making engagement in the chat superfluous: "I didn't remember to talk with my partner, because I just want to write in my own way. Because she was mainly correcting my writing, so I think she can check whether it is right or wrong."

These findings show that the online environment of collaborative writing tools such as Google Docs affects the dynamics of writing (Godwin-Jones, 2018). In contrast to offline word processor tools, the simultaneous writing and editing facility in Google Docs, potentially elicits a dominant-passive role assignment quite naturally, that is, one writes, the other corrects and edits. The result might be a text of good quality – which is often brought forward as one of the benefits of CW (Storch, 2005). Yet, the specific affordances of the online environment could have an inhibitory effect for some students, who might feel exposed while writing. Those that do not feel comfortable with the idea of having somebody look over their shoulder as they write might deliberately choose the role of editor, leading to less participation. For CSCW to be successful in providing L2 learning advantages (Storch, 2016), it is imperative to establish how tasks should be designed, how pairs should be formed, and what communication modes should be encouraged in order to elicit fruitful interaction that potentially benefits SLA.

In this sense, it must be noted that, in the current study, participants could only interact with each other via text chat. In general, pairs engaged in relatively short conversations, resulting in a limited amount of interaction. Earlier research suggests that learners interacting face-to-face or via voice-chat communicate more and also engage in more instances of negotiation of meaning than those using text chat (Cho, 2017; Rouhshad & Storch, 2016). In other words, the written mode of interaction potentially directly affected the collaborative patterns, for example, because students perceived text chat as too laborious. On the other hand, some learners might feel more comfortable contributing to written chat given that its slower pace and permanence on the screen leaves them with more time and resources for formulation, which in turn could lower anxiety levels (Satar & Özdener, 2008).

Computer-supported collaborative writing as observational learning tool (Pair 4 – expert/novice)

Pair 4 displayed a clear pattern of *expert* and *novice*. Expert Arjun produced the largest share of the final text and edits, creating the most unbalanced equality of participation score (i.e., 76.6) of the four pairs, even though Kazuo (novice) did make some edits to Arjun's text. Arjun also dominated the chat interaction (cf., Excerpt 2). While Kazuo accepted the role of novice admitting to not understand

the wording (time stamp 10:16), he also initiated a couple of content suggestions (e.g., time stamp 10:19). But Arjun disagreed, presented his own ideas and Kazuo did not pursue his own stance, adopting a more passive role. At the end (time stamp 10:40), Arjun actively encouraged Kazuo to give him feedback, but at the same time reinforced his dominant role as author by stating "if **i** need to change or add anything" (emphasis added).

Excerpt 2. Chat interaction Pair 4

Time stamp	Part.	Text:
[10:15]	Arjun:	so .. we need to address max as a student abroad advisor:
[10:16]	Kazuo:	sorry I do not understand what "address" means.:
[10:17]	Arjun:	okk .. we need to give him advice and options [**stating**]:
[10:17]	Kazuo:	Yes that is right. [**agreeing**]:
[10:18]	Arjun:	and the options will be based on the three important factors ….: cost, living and quality of studies [elaborating]:
[10:19]	Kazuo:	yes. All things considered, Canada is suitable for him, isn't it? What do you think? [**agreeing, suggesting**]:
[10:20]	Arjun:	may be … bur I think we should not give him a final solution… [**disagreeing**] we should tell him which country is affordable … which ountry is better to live .. and which country will provide hime a good education [**suggesting**]:
[10:22]	Kazuo:	good idea! I agree [**agreeing**]:
[10:27]	Kazuo:	How about writing about UK before the USA? UK is top 1 among all topic. After USA, will you write about that? [**suggesting**]:
[10:27]	Arjun:	.. i think we will first present the mediocire options like usa and australia …. them will present UK which is undoubtedly beter than all [**disagreeing**] as it will then create a bigger impact:
[…]:		
[10:40]	Arjun:	let's read it again, and pls suggest me if i need to change or add anything: [**requesting**]

The eye-gaze data mirror these roles. Arjun focused to a large extent on the writing pane, writing his own text, while Kazuo frequently checked the task graphics and Arjun's chat messages, but also spent substantial amount of time looking at the writing pane, presumably, monitoring the emerging text as editor. As such, he demonstrated constant engagement through reading.

The stimulated recall interviews reveal that both students were aware of their own and their partner's roles. Arjun noticed "For a certain time, I felt that I might be dominating this writing part, so I just tried to make Kazuo comfortable and

make sure that he has the right to interrupt me […]." Adding later that Kazuo is "someone who is a little passive". In alignment, Kazuo identified Arjun as: "the leading part, so I should respect his ideas. […] that's why I wanted to follow his ideas […]." Similar to Jingmin, he accepts that "My part was kind of spellchecker […]. I tried to check his ideas."

This last comment highlights an important aspect of CSCW that potentially supports language learning: The online environment for CW makes it possible for novices to observe the writing process and emerging text of a (more expert) partner in real-time. Accordingly, it may function as a context for observational learning (i.e. learning from models). This has been shown to promote students' writing skills given that it allows them to focus and reflect on the writing approach of the model, potentially increasing students' knowledge about writing (Braaksma et al., 2004). Previous research on text-based CMC argues that its increased salience is particularly supportive of L2 learning (Sauro, 2009; Smith, 2005). In line with Michel and O'Rourke (2019) and Michel and Smith (2018), our data show that the use of eye-tracking combined with stimulated recall interviews seems to be particularly valuable in determining which parts of the online text is noticed by a learner. Future work in the field of CSCW can explore how these instances of noticing relate to a participant's subsequent writing and editing behaviour and, in the longer run, to L2 learning.

Conclusion

This study explored L2 learners' patterns of interaction during CSCW using Google Docs triangulating data from text mining, analyses of chat logs, eye-tracking, and stimulated recall interviews to identify the levels of equality and mutuality within EFL learning pairs.

Our findings illustrate the complexities of interaction in CSCW, as participants demonstrated considerable variation regarding equality and engagement with the input of the partner through peer editing, chat interaction, and attention to/for the developing text and their partner's chat messages. Importantly, our data suggest that CSCW can contribute to L2 development given that learners can observe how a partner writes a text. This seems to be particularly fruitful in an expert/novice setting. However, the online environment, and the fact that interaction was only possible via written chat, also increased the chance that some learners adopted a passive role. By reducing themselves to 'spell-checkers', they easily disengaged from the task of jointly writing a text, which restricted their opportunities to learn from CSCW. To a large extent, contribution, engagement, and, thus, learning seems to

depend on the partner. Practitioners wishing to employ CSCW in the language classroom will need to think carefully about how to pair students, and how to design tasks that foster collaborative and/or observational learning but counter passive disengagement.

From a methodological point of view, we have shown that triangulating data from several data sources gives us rich insights into the different patterns of interaction during CSCW. In particular, the stimulated recall interviews allowed for a more valid classification of the pairs, while the eye-gaze data provided information on less salient forms of engagement in the task, for example, through reading, without direct interaction (i.e., negotiation of meaning, editing). Consequently, our findings postulate that the language learning potential of CSCW might not be restricted to learners that actively contribute to collaborative interaction. Finally, our study reiterates the call for a more dynamic approach to classifying pair interaction, as participants showed changing patterns throughout the tasks.

Given its exploratory nature, the current study has some limitations, which we will discuss alongside some methodological insights that might inform future work. First, analysing eye-gaze data in a dynamic context, such as text production in Google Docs, holds many challenges because all elements in CSCW are dynamic and interactive (Michel & Smith, 2017). For example, participants could scroll up and down between the source graphics and the evolving text. Similarly, the chat contributions moved upwards with every new message added to the conversation. In contrast to highly controlled psycholinguistic experiments, marking the Areas of Interest in our data required manual second-by-second coding (Michel & Smith, 2018). In addition, practicalities made us use two different systems to collect the eye-gaze data. Those constraints might have introduced some measurement inaccuracies in our study.

Second, students participated in the stimulated recall interviews one after the other, which means that for half of the participants, there was a 45 minute break between task performance and the interview. It might be that the ability to recall their thoughts during the task was different for those students with a shorter/longer wait.

Third, it was beyond the scope of the current study to look at quality of the jointly written texts. For future work, it would be interesting, though, to explore how the different interactional patterns in CSCW relate to the writing product, presumably, shedding more light on the language learning potential of CSCW.

To conclude, our research confirmed that CSCW affords a wide range of interactional patterns, which provide various opportunities for language learning. Even though only one pair displayed a collaborative pattern, all participants demonstrated engagement with the writing task and the writing process, often simultaneously through observing, editing, or discussion with the partner. Therefore, the

study of CSCW as a site for language learning should feature prominently in future research agendas on the connection between L2 writing and L2 learning. It is hoped that methodological insights provided in our contribution to this book can inform future work in the domain.

References

Abrams, Z. (2016). Exploring collaboratively written L2 texts among first-year learners of German in Google Docs. *Computer Assisted Language Learning*, 29(8), 1259–1270. https://doi.org/10.1080/09588221.2016.1270968

Arnold, N., Ducate, L., & Kost, C. (2012). Collaboration or cooperation? Analyzing group dynamics and revision processes in wikis. *CALICO Journal*, 29(3), 431–448. https://doi.org/10.11139/cj.29.3.431-448

Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23). New York, NY: Guilford Press.

Braaksma, M. A. H., Rijlaarsdam, G., van den Bergh, H., & van Hout-Wolters, B. H. A. M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition and Instruction*, 22(1), 1–36. https://doi.org/10.1207/s1532690Xci2201_1

Cho, H. (2017). Synchronous web-based collaborative writing: Factors mediating interaction among second-language writers. *Journal of Second Language Writing*, 36, 37–51. https://doi.org/10.1016/j.jslw.2017.05.013

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research*, 32(3), 453–467. https://doi.org/10.1177/0267658316637401

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7(4), 482–511. https://doi.org/10.1177/0741088390007004003

Damon, W., & Phelps, E. (1989). Critical distinction among three approaches to peer education. *International Journal of Educational Research*, 58(2), 9–19. https://doi.org/10.1016/0883-0355(89)90013-X

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian apporaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.

Eddy-U, M. (2015). Motivation for participation or non-participation in group tasks: A dynamic systems model of task-situated willingness to communicate. *System*, 50, 43–55. https://doi.org/10.1016/j.system.2015.03.005

Fernández Dobao, A., & Blum, A. (2013). Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System*, 41(2), 365–378. https://doi.org/10.1016/j.system.2013.02.002

Godwin-Jones, R. (2018). Second language writing online: An update. *Language Learning & Technology*, 22(1), 1–15.

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11(4), 329–350. https://doi.org/10.1016/S1060-3743(02)00091-7

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, 24(04), 541–577. https://doi.org/10.1017/S0272263102004023

Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. *Language Learning & Technology*, 16(1), 91–109.

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, 12(2), 211–234. https://doi.org/10.1177/1362168807086288

Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning*, 20(3), 1–23. https://doi.org/10.1080/09588221.2018.1465981

Li, M., & Kim, D. (2016). One wiki, two groups: Dynamic interactions across ESL collaborative writing tasks. *Journal of Second Language Writing*, 31, 25–42. https://doi.org/10.1016/j.jslw.2016.01.002

Li, M., & Zhu, W. (2013). Patterns of computer-mediated interaction in small writing groups using wikis. *Computer Assisted Language Learning*, 26(1), 61–82. https://doi.org/10.1080/09588221.2011.631142

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bathia (Eds.), *Second language acquisition, Vol. 2: Handbook of language acquisition* (pp. 413–468). New York, NY: Academic Press.

Mackey, A. (2007). The role of conversational interaction in second language acquisition. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 1–26). Oxford: OUP.

Manchón, R. M. (2011). Situating the learning-to-write and writing-to-learn dimensions of L2 writing. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 3–16). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.03man

Manchón, R. M., & Roca de Larios, J. (2007). On the temporal nature of planning in L1 and L2 composing. *Language Learning*, 57(4), 549–593. https://doi.org/10.1111/j.1467-9922.2007.00428.x

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Michel, M., & O'Rourke, B. (2019). What drives alignment during text chat with a peer vs. a tutor? Insights from cued interviews and eye-tracking. *System*, 83(50–63). https://doi.org/10.1016/j.system.2019.02.009

Michel, M., & Smith, B. (2017). Eye-tracking research in computer-mediated language learning. In S. May (Ed.), *Language and technology: Encyclopedia of language and education* (pp. 1–12). Dordrecht: Springer.

Michel, M., & Smith, B. (2018). Measuring lexical alignment during L2 chat interaction: An eye-tracking study. In S. M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition*. London: Routledge.

O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer mediated communication and language learning. *Computer Assisted Language Learning*, 21(3), 227–251. https://doi.org/10.1080/09588220802090253

O'Rourke, B. (2012). Using eye-tracking to investigate gaze behaviour in synchronous computer-mediated communication for language learning. In M. Dooly & R. O'Dowd (Eds.), *Researching online interaction and exchange in foreign language education: Methods and issues* (pp. 305–341). Frankfurt: Peter Lang.

Polio, C. (2012). The acquisition in of second language writing. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 319–334). London: Routledge.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction* (pp. 211–219). Hershey, PA: Idea Group Reference. https://doi.org/10.4018/978-1-59140-562-7.ch034

Reichle, E. D. (2006). Computational models of eye-movement control during reading: Theories of the "eye–mind" link. *Cognitive Systems Research*, 7(1), 2–3. https://doi.org/10.1016/j.cogsys.2005.07.001

Révész, A. & Michel, M. (Eds.) (2019). *Methodological advances in investigating L2 writing processes*. Special Issue, *Studies in Second Language Acquisition*, 41(3).

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors: A mixed methods study. *Studies in Second Language Acquisition*, 41, 605–631. https://doi.org/10.1017/S027226311900024X

Roca de Larios, J., Manchón, R. M., & Murphy, L. (2006). Generating text in native and foreign language writing: A temporal analysis of problem solving formulation processes. *Modern Language Journal*, 90(1), 100–114. https://doi.org/10.1111/j.1540-4781.2006.00387.x

Roca de Larios, J., Nicolás-Conesa, F., & Coyle, Y. (2016). Focus on writers: Processes and strategies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 267–286). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-015

Rouhshad, A., & Storch, N. (2016). A focus on mode: Patterns of interaction in face-to-face and computer-mediated contexts. In M. Sato & S. G. Ballinger (Eds.), *Language learning & language teaching. Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 267–289). Amsterdam: John Benjamins.

Rouhshad, A., Wigglesworth, G., & Storch, N. (2016). The nature of negotiations in face-to-face versus computer-mediated communication in pair interactions. *Language Teaching Research*, 20(4), 514–534. https://doi.org/10.1177/1362168815584455

Satar, M., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *Modern Language Journal*, 92(4), 595–613. https://doi.org/10.1111/j.1540-4781.2008.00789.x

Sauro, S. (2009). Computer-mediated corrective feedback and the development of second language grammar. *Language Learning & Technology*, 13(1), 96–120.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: CUP. https://doi.org/10.1017/CBO9781139524780.003

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly*, 39(1), 33–58. https://doi.org/10.2307/3588451

Smith, B. (2009). The relationship between scrolling, negotiation, and self-initiated self-repair in an SCMC environment. *CALICO Journal*, 26(2), 231–245. https://doi.org/10.1558/cj.v26i2.231-245

Smith, B., & Renaud, C. (2013). Using eye tracking as a measure of foreign language learners' noticing of recasts during computer-mediated writing conferences. In K. McDonough & A. Mackey (Eds.), *Interaction in diverse educational settings* (pp. 147–166). Amsterdam: John Benjamins.

Stickler, U., & Shi, L. (2015). Eye movements of online Chinese learners. *CALICO Journal*, 32(3), 380–395. https://doi.org/10.1558/cj.v32i3.27737

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research*, 5(1), 29–53. https://doi.org/10.1177/13621688010050103

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158. https://doi.org/10.1111/1467-9922.00179

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153–173. https://doi.org/10.1016/j.jslw.2005.05.002

Storch, N. (2009). *The nature of pair interaction. Learners' interaction in an ESL class: its nature and impact on grammatical development*. Saarbrücken, Germany: VDM Verlag.

Storch, N. (2011). Collaborative writing in L2 contexts: Processes, outcomes, and future directions. *Annual Review of Applied Linguistics*, 31, 275–288. https://doi.org/10.1017/S0267190511000079

Storch, N. (2013). *Collaborative writing in L2 classrooms. New perspectives on language and education*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847699954

Storch, N. (2016). Collaborative writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (387–406). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-021

Storch, N. (2017). Implementing and assessing collaborative writing activities in EAP classes. In J. Bitchener, N. Storch, & R. Wette (Eds.), *Teaching writing for academic purposes to multilingual students: Instructional approaches* (pp. 130–142). New York, NY: Routledge. https://doi.org/10.4324/9781315269665-9

Storch, N., & Aldosari, A. (2013). Pairing learners in pair work activity. *Language Teaching Research*, 17(1), 31–48. https://doi.org/10.1177/1362168812457530

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (97–114). Oxford: OUP.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contributions of Halliday and Vygotsky* (pp. 95–108). London: Continuum.

Swales, J., & Feak, C. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd). Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.2173936

Tan, L. L., Wigglesworth, G., & Storch, N. (2010). Pair interactions and mode of communication: Comparing face-to-face and computer mediated communication. *Australian Review of Applied Linguistics*, 33(3), 27.1–27.24. https://doi.org/10.2104/aral1027

Wang, D. (2016). DocuViz. https://doi.org/10.1145/2702123.2702517

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121–142. https://doi.org/10.1177/136216880607074599

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445–466. https://doi.org/10.1177/0265532209104670

Yim, S., Wang, D., Olson, J., Vu, V., & Warschauer, M. (2017). Synchronous collaborative writing in the classroom: Undergraduates' collaboration practices and their impact on text quality, quantity, and style. *Proceedings of the 20th ACM conference on computer-supported cooperative work and social computing* (pp. 468–479). https://doi.org/10.1145/2998181.2998356

Yim, S., & Warschauer, M. (2017). Web-based collaborative writing in L2 contexts: Methodological insights from text mining. *Language Learning & Technology*, 21(1), 146–165.

Young, A., & Tedick, D. J. (2016). Collaborative dialogue in a two-way Spanish/English immersion classroom. In M. Sato & S. G. Ballinger (Eds.), *Language learning & language teaching. Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 135–160). Amsterdam: John Benjamins.

The meaning-making potential of collaborative L2 writing at tertiary level

Marcus Saller

University of Paderborn

This chapter reports on a longitudinal study on collaborative writing by advanced L2 university students. In an attempt to elucidate the language-learning potential of collaborative writing over one semester, a mixed-method, descriptive, exploratory study with students writing essays in digital environments individually ($n = 11$) and in pairs ($n = 15$) was designed. Data consisted of the audio recordings of dyadic interaction, surveys, and expository essays. Results provide evidence of dyads deliberating upon complex meaning-making decisions, indicate positive responses to collaborative writing, and suggest heterogeneous development of syntactic complexity features characteristic of academic writing between the groups. Findings will be discussed from the point of view of research and pedagogy.

Introduction

Collaborative L2 writing as a site for language development and its purported superiority over individual writing is grounded in cognitive and sociocultural theories of second language acquisition (SLA). The cognitive rationale rests on the assumption that combined oral and written language production promotes a stronger focus on form, thought to be conducive to language learning. From a sociocultural angle, collaborative writing is assumed to be superior to individual writing because it enables learners to co-construct new L2 knowledge by pooling their linguistics resources. However, even in combination, these perspectives still appear insufficient to capture the full language learning potential of collaborative writing as a process of joint meaning-making. As noted by Byrnes and Manchón (2014), "in the act of composing … writers re-semioticize existing realities and create entirely new worlds of meaning. If, then, writing is fundamentally about how we go about making meaning, it is at the same time a way for writers to participate in constructing new knowledge" (p. 6). Accordingly, it has been suggested that

future research agendas on writing as a site for language learning ought to include a textual meaning-making orientation.

The research reported in this chapter is an attempt to move in this direction. The study explores the meaning-making potential of collaborative L2 writing at tertiary level and the development of syntactic complexity in expository essays of advanced EFL learners. I begin by offering a theoretical framework for my study, grounded in cognitive, sociocultural, and functional theories of SLA, coupled with a review of relevant research on L2 writing and syntactic complexity.

Theoretical framework

Cognitive perspective

The psycholinguistic rationale for the language learning potential of collaborative L2 writing derives from cognitive SLA models, most notably Long's (1996) Interaction Hypothesis and Swain's (1985, 1995) Output Hypothesis. According to Long (1996), interactional adjustments triggered by interlocutors' implicit form-focused meaning negotiations and negative feedback during interaction drive interlanguage development. The central tenet of the Output Hypothesis postulates that language production, in general, facilitates L2 learning when students are 'pushed' to attend to form-meaning encodings in order to meet communicative goals. Compared to spoken output, the lower online processing demands of writing and its permanence potentially allow L2 learners to allocate more attentional resources to formal language dimensions and notice gaps in their language repertoires (Williams, 2012).

Communicatively challenging writing is viewed as a cognitively complex activity that involves continuous problem solving (Manchón & Roca de Larios, 2007). From this perspective, advanced individual L2 writers self-initiate attention to form and may engage in deep problem-solving behavior in their attempts to bridge gaps between content and rhetorical concerns, a process understood to be particularly favourable for language development. As argued by Manchón and Williams (2016), "the deeper linguistic processing associated with the meaning-making activity that characterizes complex forms of writing will prompt L2 users to engage in crucial language learning processes, such as noticing, or metalinguistic reflection/analysis of explicit knowledge" (p. 572).

Cumming's (1990) empirical account of L2 writers' dynamic meaning-making behaviour during text-generation demonstrated the instrumental role that demanding composition writing may play in L2 learning, which he described as follows:

> Composing might function broadly as a psycholinguistic output condition wherein learners analyze and consolidate second language knowledge that they have previously (but not yet fully) acquired … [and] elicits an attention to form-meaning relations that may prompt learners to refine their linguistic expression – and hence their control over their linguistic knowledge …   (p. 483)

Cumming's analysis of think-aloud protocols revealed that approximately 30% of decision-making episodes in his data involved ideational and metalinguistic concerns in tandem, "show[ing] potential value for incidental learning of the second language" (p. 482). Cumming found his participants exerting cognitive effort to search their mental lexicon for words and phrases that matched their communicative intent, encompassing three types of cognitive strategies: (1) searching out and assessing improved phrasing; (2) comparing cross-linguistic equivalents; and (3) reasoning about linguistic choices.

Collaborative writing may conjure up even more language learning opportunities than individual writing for various reasons. First, during meaning negotiations, dyads (or small groups) may notice more language gaps than students writing individually, and interactional negotiation strategies (e.g. clarification requests, confirmations checks) may lead to deeper levels of language awareness. Second, learners writing collaboratively can draw on their combined explicit L2 knowledge to resolve their communication problems. Third, feedback for output modification is not delayed but immediate and continuous. Fourth, the L2 is used instrumentally in form and meaning negotiations, further elaborated by Swain's (2000) concept of collaborative dialogue, discussed in more detail below.

Sociocultural perspective

The sociocultural rationale for the language learning potential of collaborative writing draws on the later work of Swain (2000, 2006). Swain's (2000) notion of *collaborative dialogue* is conceptualised as a cognitive tool that mediates problem-solving behaviour and language development. Collaborative dialogue has been analysed in terms of *language-related episodes* (LREs), which are instances of joint written (or oral) text production in which learners articulate their thoughts or deliberations upon form and/or meaning (Swain & Lapkin, 1995). Swain (2006) termed these episodes *languaging* i.e., "the process of making meaning and shaping knowledge and experience through language" (p. 98).

In individual writing, languaging manifests itself in private speech; yet individual writers are left to their own devices when they notice linguistic problems. Collaborative dialogue, in contrast, is shaped by the synergy of students'

externalised private speech and thus may promote a deeper level of awareness of the relationship between "meaning, form and function in context" (Swain & Lapkin, 1998, p. 330). In the act of joint text production, dyads can build upon each other's deliberations involving interactional moves and collectively scaffold each other's performance by pooling their linguistic resources (Donato, 1994).

These predictions have received empirical confirmation. For instance, Kim's (2008) longitudinal study compared the impact of collaborative dialogue and private speech on L2 vocabulary acquisition. Kim found that the students working in pairs resolved more LREs correctly on a pre-test (dictogloss) and also scored significantly better individual results on post-tests than the students working alone. Similarly, Storch (2008) examined the quality of metatalk in terms of dyads' level of engagement on a text-reconstruction task. Elaborate engagement, featuring instances of collective scaffolding, resulted in a higher proportion of correctly resolved LREs and positively influenced individual performance on a post-test.

Despite these findings, research on the quality of dyadic metatalk in complex meaning-making environments appears to be unchartered territory. Similarly, longitudinal research conducted with advanced EFL learners is scarce, and only a few studies have reported on L2 learners' generally positive perceptions of collaborative writing (e.g. Shehadeh, 2011; Storch, 2005). Yet, this line of inquiry may further our understanding of learners' meaning-making behaviour since these perceptions strongly influence the effectiveness of collaborative writing tasks.

A small number of studies have compared individual with collaborative writing task performance. By and large, dyads tend to obtain higher accuracy scores in form-focused writing activities (e.g. Reinders, 2009; Storch, 1999) and produce more accurate compositions than students writing alone (e.g. Storch, 1999, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). For example, Storch and Wigglesworth (2007) and Wigglesworth and Storch (2009) compared the performance of advanced ESL students writing argumentative essays individually and collaboratively. Both studies report a strong focus on lexical deliberations during collaborative engagement, instances of collective scaffolding, and more accurate essays by dyads. Yet, given the field's preoccupation with the analysis of linguistic accuracy, little is known about the impact of peer interaction on syntactic complexity as an integral part of academic language development, which is a focal point of the study reported in this chapter, as explained in the next section.

Functional perspective

Syntactic complexity broadly refers to form range and sophistication in language production. The construct has been widely used in SLA-oriented writing scholarship to determine linguistic development. According to Ortega (2015), "it is posited that syntactic complexity indexes the expansion of the capacity to use the additional language in ever more mature and skillful ways … to fulfill various communicative goals successfully" (p. 82). In light of the above, a focused examination of syntactic complexity alone was chosen to determine the connection between collaborative writing and L2 students' progress in their academic writing.

The operationalisation of syntactic complexity has presented an empirical challenge, as shown by the variety of proposed developmental indices (e.g. Biber, Gray, & Poonpon, 2011; Larsen-Freeman, 2006; Norrby & Håkansson, 2007). What is more, corpus-based research by Biber and colleagues (e.g. Biber & Gray, 2016) has demonstrated that commonly employed subordination measures are insufficient for gauging complexity in L2 academic writing. Clausal subordination is appropriate for determining complexity in speech whereas phrasal complexity is a suitable measure for academic writing.

In recent years, L2 writing scholarship has approached syntactic complexity through a functional lens to unravel "the meaning dimension of complexity" and account for the communicative demands imposed by academic writing tasks (Ryshina-Pankova, 2015, p. 52). The research reported in this chapter follows suit and draws on Systemic Functional Linguistics (SFL), which is sensitive to L2 development and re-frames grammar as a meaning-making resource (Halliday & Matthiessen, 2004). SFL identifies grammatical metaphor (GM) as a distinguishing characteristic of written discourse (Halliday, 2002), and the construal of GM in the *ideational domain*, as Byrnes (2014) argues, "constitutes a critical step in a language user's cognitive and linguistic development" (p. 96). Ideational GM is often realised through the nominalisation of verbs or adjectives enabling writers to create and foreground abstraction and expand meaning through noun-phrase modifications (Byrnes, 2009; Liardét, 2013). Longitudinal findings provide evidence that advanced L2 student writers increasingly employ ideational GM in order to meet the challenges of complex academic writing tasks (e.g. Byrnes, 2009; Whittaker & McCabe, Chapter 13, this volume). However, to date, research has been confined to individual writing conditions.

The present study operationalises Biber et al.'s (2011) developmental index for a targeted analysis of syntactic complexity in L2 academic writing. The five-stage index progresses along formal and functional parameters from clausal to phrasal complexity and predicts the acquisition of academic complexity features, such as

complex pre- and postmodifiers of noun phrases, at stages four and five. Yet, thus far, empirical evidence, particularly from longitudinal research, is scarce.

Parkinson and Musgrave (2014) narrowed the analytical scope of the index to noun-phrase complexity and compared academic compositions of two L2 proficiency groups. Results include a significantly greater proportion of prepositional phrases with heads other than 'of' and abstract meaning (stage four) in the writing of advanced students. Mazgutova and Kormos (2015) examined syntactic complexity development in argumentative essays over a four-week period and found a significant increase of noun-phrase complexity in the intermediate group whereas this feature marginally decreased in the upper-intermediate group. The researchers suggest that the development of abstract lexical sophistication may occur at the expense of syntactic complexity in advanced L2 writing. Longitudinal growth in noun-phrase complexity is also reported in Crossley and McNamara's (2014) study of L2 descriptive essays over a semester-long writing course.

The present study

Aims and research questions

With the ultimate aim of elucidating the connection between collaborative L2 writing and language learning, the goals of the present study were to explore complex meaning-making behaviour in dyads, students' attitudes towards collaborative writing, and the emergence of syntactic complexity as a result of collaborative writing at L2 tertiary level. The study was guided by three research questions:

1. Does collaborative L2 writing push dyads to engage in complex meaning negotiations?
2. How do students evaluate collaborative L2 writing at tertiary level?
3. Does collaborative L2 writing promote the development of syntactic complexity features characteristic of academic writing?

Method

Research site

The study was conducted in the English Language Teaching program at a large university in northern Germany. Data were collected during the winter term of 2017/18 in two parallel "Collaborative Writing" seminar courses taught by the researcher. Seminar attendance was not compulsory.

Research approach

The study employed a mixed-method, descriptive, and exploratory approach to shed light on the meaning-making potential of collaborative L2 writing at tertiary level making use of qualitative and quantitative data collected at different times (see Table 1).

Table 1. Distribution of data collection

Time	Expository essay	Audio recording	Survey 1	Survey 2
1	✓		✓	
2	✓	✓		✓
3	✓	✓		✓
4	✓	✓		✓
5	✓	✓		✓
6	✓	✓		✓
7	✓	✓		✓
8	✓	✓		✓
9	✓	✓		✓
10	✓		✓	

Students in both groups (henceforth, Group 1 and Group 2) wrote expository essays about the same topics (see Appendix A). At Time 1 and Time 10, all essays were handwritten by individual students to avoid technology dependence. From Time 2 to Time 9, all participants composed their essays on Google Docs, using their own laptops: individually in Group 1; in pairs in Group 2[1]. Computers were considered the more appropriate medium for this phase of the study than pen and paper since students in Group 2 could easily decode unfolding text on their computer screens and switch back and forth seamlessly between reader and writer roles.

Participants

Twenty-six first-year Master of Education (M.Ed.) students were invited to participate in this study. They all signed a consent form. Both Group 1 ($n = 11$) and Group 2 ($n = 15$) were composed of monolingual native German speakers, predominantly female, with respective mean ages of 23.5 and 25.1 years. All participants had been exposed to English as the medium of instruction at university for between three and four and a half years and had practiced expository essay writing in the

---

1.   Data collected from groups of three students were not included in the analysis.

compulsory undergraduate course "Academic Skills". The participants' academic writing proficiency was assessed with the standard IELTS rubric for academic writing. Group A outperformed Group B by approximately half an IELTS band score with a mean of 6.2 compared to 5.8 (see Table 2).

Table 2. Sample characteristics

	Group 1 ($n = 11$)	Group 2 ($n = 15$)
L1 (L2)	German (English)	German (English)
Mean age	23.5	25.1
Female	11	10
Male		5
Mean IELTS writing	6.2 (B2 on the CEFR*)	5.8 (B2 on the CEFR*)

* Common European Framework of Reference for Languages (CEFR)

Instruments

Data for this study were obtained from three different sources: audio recordings, student surveys and expository essays. Data from audio recordings were analysed to address Research Question 1 ('Does collaborative L2 writing push dyads to engage in complex meaning negotiations?'), data from student surveys for Research Question 2 ('How do students evaluate collaborative L2 writing at tertiary level?'), and data from expository essays for Research Question 3 ('Does collaborative L2 writing promote the development of syntactic complexity features characteristic of academic writing?').

Audio recordings

The first main data source for this study came from audio recordings of pair talk. Prior to each collaborative writing session (Times 2–9), work stations with digital recording equipment had been installed in the classroom. Since attendance fluctuated during this period, participants were assigned to their work stations *ad hoc* in order to rotate to a new partner every week. Thus, data were obtained for the analysis of dyadic interaction and a comparison of individual behaviour in different dyadic constellations. The recording equipment was synchronised with Google Docs, which creates timestamps at short intervals, allowing the reconstruction of the stages of the writing process, which could then be directly related to the corresponding episodes of transcribed discourse.

Surveys

Surveys constituted the second data source. All survey items were created by the author and his colleagues because despite extensive searches, an appropriate survey for this study could not be found in relevant SLA literature. Two types of surveys were designed, piloted and revised. The main survey included 50 belief statements about pair work and L2 academic writing on a five-point Likert scale, open-ended questions eliciting additional information, and a section with background questions. A scale of 21 closed-ended items explored respondents' perceptions of collaborative writing as fostering the development of academic writing skills (see Appendix B). This survey was completed by students from Group 2 at Time 1 and Time 10. In addition, from Time 2 to Time 9, participants in both groups completed a short retrospective survey after each writing activity, which asked about their experience during the writing process (see Appendix C).

Expository essays

Expository essays were chosen for a focused examination of syntactic changes in the ideational domain, since this genre favours impersonal register and grammatical metaphor (Martin, 1989). All essays were written under a 60-minute time constraint; the use of dictionaries was not allowed. In order to stimulate the quality of written output, two identical essay prompts suggesting topic familiarity were used at Time 1 and Time 10: *Write an academic essay to discuss the significance of education on life quality*. However, task repetition may have influenced the students' writing performance at Time 10 and should therefore be considered a potential drawback of this study. From Time 2 to Time 9, the essay topics were determined by preceding student presentations (see Appendix A).

Data analysis procedures

Discourse analysis of dyadic interaction

The first data set for analysis derived from 29 hours of recorded pair talk in 29 unique constellations. The method of conversation analysis was chosen for a fine-grained examination of sequentially unfolding meaning negotiations and the dynamics of dyadic interaction. Pseudonyms were used to preserve the participants' anonymity.

The present study examined the students' negotiations of topic sentences and thesis statements (*n = 96*) as vehicles of progressive meaning-making in expository essays. Eighty percent of corresponding discourse was transcribed by one of the author's M.Ed. students, who would later conduct an independent analysis (Göknil, 2018). The remaining transcriptions were completed, and all transcriptions were revised by the author, comprising 1,883 turns, which were then segmented into *language-related episodes* and *non-linguistic episodes*.

Language-related episodes

Language-related episodes (LREs) were identified as "any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others" (Swain & Lapkin, 1998, p. 326). LREs were coded as lexis-focused (i.e. meaning, word choice) and form-focused (i.e. grammar, mechanics).

The initial coding process revealed dyads primarily deliberating upon better language choices for their essays rather than tackling linguistic or semantic problems because of limited L2 repertoires. In order to explore textual meaning-making in an appropriate manner and identify and discriminate between episodes indicating language limitations and episodes indicating higher levels of ideational concerns, the meaning dimension of the LRE taxonomy was relabelled 'lexicogrammar' under an SFL-optic and configured into two broad categories, namely, *lexicogrammar problem* and *lexicogrammar alternative*.

The first category, *lexicogrammar problem* (LGP), comprised transcribed segments of dyads either assessing the meaning and appropriateness of L2 vocabulary, or "making cross-linguistic comparisons" (Cumming, 1990, p. 493), as illustrated in episode (1).

```
   (1)  LGP L1–L2
1   Asta   →   was heißt nochmal entlehnt? (2.0)
                how do you say 'derive' again?
2   Ela    →   hm?=
3   Asta   →   =like the concept is taken from (2.0) um: the early um:
                (2.0) stages of language
4               learning (.) in which a caretaker (.) like >for instance
                the mother<=
5   Ela        =okay=
```

In episode (1), Asta asks Ela for the translation of 'derive' (line 1). Ela's clarification request (line 2) pushes Asta to formulate the paraphrase 'taken from' (line 3), which is appropriate for the target sentence.

The second category, *lexicogrammar alternative* (LGA), was further segmented
into 'word', 'phrase', 'clause' and 'nominalisation' to capture deliberations along the
lexicogrammar stratum and examine complex meaning-making processes, illus-
trated in episodes (2), (3), (4) and (5), respectively.

(2) LGA word

```
30   Ina   →   =interaktion ((pronounced like German equivalent))
                 it is helpful to visualise (.) or to explain
31         →   whatever you=
32   Ute       =to (1.0) ne ek yeah: (1.0) geht beides ne?=
                                         both work, no?
33   Ina       =yeah=
34   Ute   →   =((reads off screen)) it is helpful to (1.0) or
                 introduce or=
35   Ina   →   =introduce the model of dyadic interaction
```

In episode (2), two students negotiate an appropriate verb of an infinitive clause in
extraposed subject position. Ina suggests 'visualise' and 'explain' (lines 30–31), but
readily accepts Ute's alternative 'introduce' (lines 34–35).

(3) LGA phrase

```
24   Ela             [into]          [a]
25   Ute       =realistic (0.5) or a (2.0)
26   Ela   →   classroom (.) learning situation=
27   Ute       =yeah >in a< (2.0)
28   Ela   →   lesson plan
29   Ute   →   lesson plan
30   Ela   →   lesson plan?=
31   Ute   →   =yeah=
```

In episode (3), Ela and Ute deliberate upon an appropriate noun phrase as a prep-
ositional complement. Ella suggests 'classroom learning situation' (line 26) and
'lesson plan' (line 28), and both students agree on choosing the latter more suitable
alternative (lines 29–31).

(4) LGA clause

```
28   Jan   →   =writers (.) we should (.) um: change the sentence
                 >a little bit<=
29   Mia       =>yeah yeah yeah<=
30   Jan   →   =because one has to consider=
31   Mia       =yeah=
32   Jan   →   ((laughs)) [isn't] (.) that good I think (.) I
                 think uh it has to be considered
33   Mia   →   yeah
34   Jan       (2.0) would be better (2.0) ((reads off screen))
                 considered (.) that the
```

In episode (4), Jan suggests revising the clause 'one has to consider' (lines 28, 30, 32) to one with impersonal passive voice (line 32). Mia agrees without objection (line 33).

(5) LGA nominalisation

```
44   Noel   →   uh: (1.0)  >let's just start< with like um: is an
                 activity that (1.5) or where (.)
45          →   learners (.) can practice (.) writing talking
                 listening (1.0) >like that< we list
46   Tom    →   or offers the opportunit[y for] learners to (1.0)
47   Noel   →                          [yeah]  that's good
48   Tom    →   ((reads off screen while typing)) offers the
                 opportunity for learners (3.5)
49          →   for learners (2.0) um: (2.0) first to communicate
                 (1.0) or to (.) yeah to
50          →   communicate about the (2.0) um: (2.0)
```

In episode (5), Noel suggests a construction with a copula and a subject complement modified by a relative clause (lines 44–45) but approves of Tom's more sophisticated proposal, namely the transitive verb 'offer' followed by a nominalised group in direct object position (lines 46–50).

LGPs and form-focused LREs were further categorised according to their outcomes as *correctly resolved* or *unresolved*. LGAs were classified as *accepted* or *rejected*.

Non-linguistic episodes

Taxonomies from Storch (2005) and Wigglesworth and Storch (2009) were adapted for the segmentation of non-linguistic episodes: *content* (idea generation and content clarification); *structure* (paragraph and essay organisation); *other* (e.g. text revision, peer dictation).

The author and a research assistant coded all transcripts independently according to the elaborated coding taxonomy (see Figure 1). Inter-rater reliability was moderate: 77% for LGPs, 72% for LGAs, 82% for form-focused LREs, and 82% for non-linguistic episodes. Discrepancies were discussed and resolved.

Survey analysis

Quantitative data obtained from the main survey, completed at Time 1 and Time 10, were entered into SPSS for descriptive statistics. Reliability measures of the scale that tapped into participants' perceptions of collaborative writing (see Appendix B) yielded high Cronbach alpha coefficients at Time 1 (α = .882) and at Time 10 (α = .887). An organic thematic analysis (Braun & Clarke, 2006) conceptualised responses to open-ended survey questions by members from both groups.
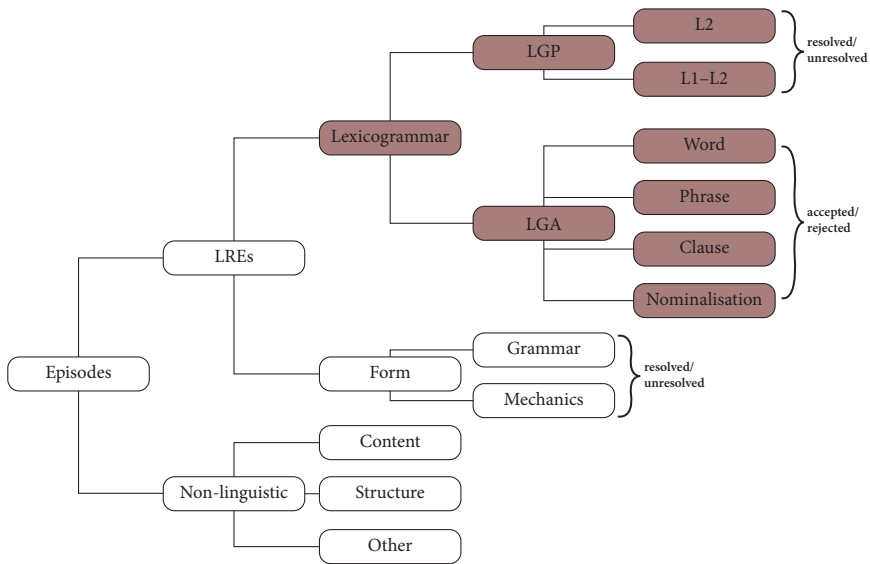
Figure 1. Coding taxonomy of this study (adapted from García Mayo & Azkarai, 2016, p. 249)

Corpus analysis of expository essays

The digitised corpus of 52 essays, handwritten by individual students in both groups at Time 1 (*n = 26*) and Time 10 (*n = 26*), was checked against the original texts by two people to ensure high levels of authenticity and then analysed in three tiers.

First, the author and a trained research assistant counted all finite and non-finite dependent clauses in the corpora manually and annotated their syntactic structures and functions in Microsoft Excel. Disagreements were discussed until resolved.

Second, syntactic errors were flagged and categorised according to their perceived gravity. Dependent clauses indicating violations of morphosyntactic rules governed by inflection (e.g. infinitival clauses after prepositions) and clauses with missing constituents (e.g. direct objects, prepositions) were removed. Grammatical structures with lexicogrammatical inadequacies or faulty punctuation (e.g. wrong subordinators, run-on sentences) were included in the analysis. Subordinate clause fragments were attached to corresponding matrix clauses for the analysis.

Third, grammatical structures, featured at stages four and five of Biber et al.'s (2011) developmental index, were analysed with non-parametric tests at a normalised frequency (f = 1,000); the UAM Corpus Tool was used to determine dependent phrases. Phrases partially or completely copied from the essay prompt ('significance of education on life quality') were removed from the analysis, and

the dependent phrases were manually re-analysed several times to aim for a high reliability. However, the absence of reliability estimates on the measures illustrated in Tables 3 and 4 should be considered a further shortcoming of this study.

Table 3. Grammatical structures at stage four of Biber et al.'s index

Structure	Example
Adj. *to*-complement clauses	Life quality depicts a vague term and seems difficult *to measure*. (SØ)
Extraposed complement clauses	
that-clauses	it is important *that this process is now guided by the teacher*.(UØ)
wh-clauses	it was already pointed out *how important it is to be educated*(OØØ)
to-clauses	it is almost impossible *to get a well paid job or even to get a job*. (GØØ)
Non-finite relative clauses	
-ing participle clauses	instances of uneducated people *finding good jobs and earning enough money* (XØ)
-ed participle clauses	opportunities *granted to a person later in life*. (HØØ)
Phrasal embedding in NP	the number of *high-school and university graduates* (NØØ)
PPs as NP postmodifiers (abstract)	
of-phrases	the process *of knowledge acquisition* (QØØ)
other than *of*-phrases	a broader perspective *on daily situations* (EØ)

Table 4. Grammatical structures at stage five of Biber et al.'s index

Structure	Example
Prep. + *ing*-complement clauses	
of-clauses	the privilege *of receiving education* (BØ)
other than *of*-clauses	the potential *for making balanced decisions in life*. (LØ)
Noun complement clauses	
that-clauses	the thesis statement *that a lack in education leads to a decrease of life quality*(RØ)
wh-clauses	the question *whether education really has an impact on life quality* (DØØ)
to-clauses	universal literacy is the goal *to achieve*. (JØ)
Appositive noun phrase	the social context of education – *the collaborative work* – improves the chance to access more social contacts (RØØ)
Ext. phrasal embedding in NP	education *as a tool in Western schools in the form of competences and skills* (AØØ)

Results and discussion

Collaborative L2 writing and meaning negotiations

Our first research question asked whether collaborative L2 writing promotes complex meaning negotiations. The quantitative analysis of the segmented and coded protocols showed that 52.6% of the episodes comprised lexicogrammatical deliberations, 9.5% were form-focused and 37.9% dealt with non-linguistic concerns, of which 58.2% tackled content issues (see Table 5). These results correspond with Wigglesworth and Storch (2009) in as much as both studies demonstrate that in open and complex composition tasks, dyads with advanced L2 proficiency primarily engage in meaning and content negotiations.

Table 5. Segmented episodes of pair talk

Focus	Episode	Sum	Resolved (%)	Accepted (%)
Lexicogrammar	Problem	24	87.5	
	– L2	1	100	
	– L1–L2	23	86.9	
	Alternative	119		86.5
	– Word	35		91.4
	– Phrase	39		87.1
	– Clause	37		78.3
	– Nominalisation	8		100
Form	Grammar	9	88.8	
	Mechanics	17	88.2	
Non-linguistic	Content	60		
	Structure	16		
	Other	27		

Meaning negotiations

In the majority of lexicogrammatical episodes (83.2%), dyads deliberated upon alternative formulations for enhanced meaning-making. Remarkably, these deliberations were evenly distributed along the lexicogrammar stratum *word-phrase-clause* (see Table 5), which arguably reflects dyads' advanced L2 proficiency. Participants were capable of formulating and assessing single words but also simple and complex phrases and clauses as more suitable alternatives for their essays. These types of meaning negotiations involved cognitively less demanding problem-solving strategies, such as "simple revisions by substitution" (Cumming, 1990, p. 492), as seen in episode (4), and cognitively challenging processes marked by longer episodes featuring long pauses between utterances, as illustrated in episode (6).

Episode (6) can be interpreted from different theoretical angles. From a combined cognitive-sociocultural perspective, Tom and Sue's elaborate engagement (Storch, 2008) and intense languaging triggers depth of processing (DoP), which, as discussed, is understood to promote language learning. Both students build upon each other's deliberations and pool their language resources to resolve a linguistic problem, resulting in the co-construction of L2 knowledge.

(6) Complex processing of target clause

```
27   Tom   →   with a lower learning outcome< (6.0) ((reads off
                screen while typing)) if the teacher
28         →   manages (3.0) to create (9.0) um: (2.0) to
29   Sue       to crea um:=
30   Tom   →   =((reads off screen while typing)) to distribute
                the class into (4.0) into uh
31         →   wie heißt es? (5.0) wenn er es schafft die klasse
                in (1.0) feste paare zu setzen die (.)
32         →   eine gute konstellation haben (2.0) dann fühlen die
                sich mehr comfOrtable
                how do you say if he manages to divide the class
                into fixed pairs which have a good
                constellation then they feel more comfortable
33   Sue       mm-hmm (2.0)
34   Tom       ((reads off screen)) >if the teacher manages to<
                [distribute ]
35   Sue   →   [then (.) or]
                something like the goal
36         →   of the teacher should be (.)
37   Tom       yeah=
```

When Tom encounters difficulties formulating a complex sentence with a concessive *if*-clause (lines 27–28, 30), he uses his L1 lexicon as a cognitive tool (lines 31–32) to attend to higher levels of discourse concerns (Cumming, 1990). At the same time, he asks his partner for help, indicated by 'wie heißt es' ('how do you say', line 31). Sue first attempts to scaffold Tom's performance by completing the *if*-clause fragment with the missing matrix clause, indicated by 'then' (line 35), however changes tack and produces an alternative syntactic plan (Roca de Larios, Murphy, & Manchón, 1999). Sue upgrades clausal complexity to phrasal complexity and the postmodified noun phrase 'the goal of the teacher' (lines 35–36) can be seen as a new variation of complex syntax in the context of academic writing, whereby form-function relationships have been restructured (McLaughlin, 1990). Similar processes have also been observed in the think-aloud data of individual L2 writers. For example, López-Serrano, Roca de Larios, and Manchón (Chapter 10, this volume) observed DoP in their advanced L2 learners searching for better language choices and "perhaps opening up new, unsuspected potential avenues of form-function mappings along the way" (p. 249).

Under a complementary SFL optic, GM dynamically emerges from Sue and Tom's joint meaning-making effort. Sue reduces Tom's complex clause construct to a single clause by metaphorically rewording 'if the teacher manages' as 'the goal of the teacher should be'. In so doing, she competently deploys abstract academic language. The abstract noun 'goal' is foregrounded as the subject of the new clause and postmodified with the prepositional phrase 'of the teacher'. This episode provides a valuable insight into the interface between writing and L2 learning in the collaborative writing condition, illustrating how dyadic interaction fosters the deployment of lexicogrammatical resources to enhance the quality of textual meaning-making (see also Byrnes, Chapter 4, this volume).

We identified seven additional episodes in which dyads negotiated GM: either by nominalising verbs, as seen in (5), where Tom metaphorically rewords 'can' (line 45) as 'opportunity' (line 46), or by nominalising adjectives, as seen in (7), where Otto metaphorically rewords 'afraid' (line 15) as 'fear' (line 20).

(7)   Nominalisation of adjective

```
15  Otto   →   because they uh: are not (0.5) ((Tine laughs))
                afraid to make mistakes (.) and
16              therefore [no no] because=
17  Tine              [yeah]
18  Otto   =the
19  Tine   teacher (.) [no ]
20  Otto   →          [tea]cher takes away the fear of making
                m mistakes? (3.0)
21  Tine   yeah=
```

In all observed instances, participants took advantage of the semogenic potential of GM and expanded meaning in noun phrases through postmodification, either with simple or multiple prepositional phrases, complementary non-finite clauses, or prepositional phrases with complementary non-finite clauses, as seen in (8).

(8)   the opportunity for learners to talk and listen to their peers in the target language   (NR2)

The findings reported in this section distinctively show that collaborative L2 writing at tertiary level promotes complex meaning negotiations potentially conducive to language learning. This, in turn, may trigger the construal of GM to "transform dynamic, grammatically intricate language into static lexically dense entities" (Liardét, 2013, p. 162). As illustrated, these deliberations may push advanced L2 learners to restructure their syntactic output during text generation in order to meet the communicative demands of challenging writing tasks, which Byrnes (2014) has called "the ability to make situated linguistic choices" (p. 87). This phenomenon remains largely undocumented in research into collaborative L2 writing, which to date has

analysed pair-talk from cognitive and sociocultural perspectives. Moreover, the emergence of GM as a key indicator of L2 development has thus far only been researched in individual writing. Future research agendas exploring the relationship between dyadic meaning-making behaviour and the linguistic and rhetorical conventions imposed by genre-oriented writing tasks in a textual environment may further elucidate the language learning potential of collaborative L2 writing.

Content negotiations

Content negotiations comprised 58.2% of non-linguistic episodes. This indicates that participants exhausted the potential of collaborative engagement as a site for idea exchange, also reported by Storch and Wigglesworth (2007). The finding correlates with the quantitative and qualitative survey results. At the beginning of the study, 60% of the participants from Group 2 agreed with the statement "I think collaborative writing with a partner would allow me to exchange ideas" and 40% strongly agreed ($M = 4.40$; $SD = .507$); at the end of the study, 73.3% strongly agreed and 26.7% agreed ($M = 4.73$; $SD = .458$). Moreover, *ideation* emerged as a positive theme from the thematic analysis, as attested by Ute in (9).

(9)   The gain of the activity was, that it could be easier to think about a topic in a pair, in order to gain even more ideas

Conversely, the analysis of qualitative survey data from Group 1 identified *ideation* as a dominant negative theme. Students frequently commented on their difficulties with generating ideas, arguments or supporting examples, as Jill in (10).

(10)   I wasn't (personally) very enthusiastic today and couldn't come up with many examples.

In Group 2, on the other hand, dyads could draw on each other's resources. In these episodes, students frequently attended to ideational and metalinguistic concerns in tandem, which has been hypothesised as facilitating language learning (Cumming, 1990).

In episode (11), Jan and Max concurrently negotiate the gist and language of a target sentence. The students pool their linguistic resources and content knowledge and collectively construct an ideational scaffold, marked by mutual confirmation and reassurance (lines 135–136, 139, 141–143, 146). In so doing, Jan and Max express a complex concept that might be beyond their individual capabilities.

(11)   Content negotiation

```
130   Jan      we have the process approach >because< that's
               what I (.) you know I just
131            said [it's]=
132   Max         [yeah]
133   Jan      =more writing a text is a little bit like uh
               (2.0) you know it's it's (1.0) to
134            craft a text (.) kind of (.)
135   Max  →   yeah it's like you you you learn the the
               craftsmanship [or  ] the the
136   Jan  →                 [yeah]
137            ((reads off screen while typing))
               process=
138   Max      =it's it's on anothe:r level it's it's on the
               meta level of [writing] (.) >you're=
139   Jan  →                 [yeah  ]
140   Max      =you're not< [just writing you]
141   Jan  →                [yeah you talk a ] bout writing
               [exactly yeah  ]
142   Max  →   [yeah >you you<] you (1.0) yeah
143        →   >you< (.) yeah I think yeah=
144   Jan      =craftsmanship (1.0) um: ach uh
               ((both laugh)) process approach um:
               (6.0)
145            takes or takes meta level into consideration
               right?=
146   Max  →   =mm-hmm
```

Instances of collective scaffolding among advanced L2 learners are also reported in Storch and Wigglesworth (2007) where advanced ESL dyads "offer suggestions and counter suggestions, seek and provide feedback to each other and build on each other's suggestions" (p. 169) while jointly composing sentences of data commentary reports. Wigglesworth and Storch (2009) describe how two advanced ESL students pool their language resources in an argumentative writing task in order to co-construct a sentence with higher levels of linguistic accuracy and complexity. The process of collective scaffolding appears to be particularly conducive to linguistic knowledge expansion or consolidation when gist and language negotiations, as illustrated in episode (11), go hand in hand.

Students' perceptions of collaborative L2 writing

Our second research question asked about our participants' evaluation of their collaborative writing experience. In what follows, we report on the quantitative and qualitative results obtained.

Quantitative survey results

At Time 1 and Time 10, the majority of respondents in Group 2 reported agreement with the belief statements about partner work in an academic writing environment (see Appendix B), with some exceptions discussed below. The result is in line with previous research on student perceptions of collaborative L2 writing, which also documented positive responses (e.g. Shehadeh, 2011; Storch, 2005).

In addition, the single item analysis revealed three noteworthy changes in student responses from Time 1 to Time 10. First, as discussed, students responded more positively to the statement concerning idea exchange at the end of the study. Second, at Time 1, 53.3% agreed with the statement "I think collaborative writing with a partner would help me to make my writing more coherent" ($M = 3.67$; $SD = .617$); at Time 10, 73.3% agreed ($M = 3.87$; $SD = .516$). Third, a positive trend was also observed with regard to the statement "I think collaborative writing with a partner would help me to make my writing more cohesive". At Time 1, 46.7% neither agreed nor disagreed and 40% agreed with the statement ($M = 3.47$; $SD = .743$); at Time 10, 66.7% agreed and 26.7% neither agreed nor disagreed ($M = 3.80$; $SD = .561$). The last two findings converge with the findings from the thematic analysis, which identified *structure* as a benefit of collaborative writing. In these responses, participants reflected upon their mostly positive experiences of structuring sentences and paragraphs with a partner, as Ina comments in (12), and thus seemed to perceive collaborative writing as a site for improving the textual and logical organisation of their essays.

(12) It was good to work with a partner in order to come up with more complex sentences and ideas.

In Group 1, *content knowledge* emerged as dominant positive theme from responses to retrospective surveys, as Nia notes in (13).

(13) I reflected on the importance of collaborative language learning and opportunities for it.

As discussed, in Group 2 content negotiations made up the greatest proportion of non-linguistic episodes. It appears that individual student writers in Group 1 compensated for the lack of partner interaction and used writing as a tool to critically reflect on theoretical input from preceding presentations, clarify comprehension issues and engender understanding.

Qualitative survey results

From Time 2 to Time 9, responses to the retrospective surveys indicate that the majority of students in Group 2 enjoyed and felt comfortable working with a partner. In line with Storch's (2002) findings, the conversation analysis identified pairs with a collaborative orientation as the dominant pattern of dyadic interaction, underscored by the high amount of resolved LGPs/form-focused LREs and accepted LGAs (see Table 5).

Conversely, the thematic analysis also revealed a negative pattern, which was determined by two factors: *time* and *intersubjectivity*. Time-management concerns were frequently addressed by respondents, as Ina remarks in (14). Oftentimes, jointly-written essays were not completed within the 60-minute time constraint. This may have caused frustration with collaborative writing, as seen in (15).

(14)   Time management complicates the production of a good essay.

(15)   All in all it was exhausting to constantly negotiate and the feeling to not get anything done was frustrating.

Another key finding from the qualitative survey analysis is participants' critical reflections on their writing partners, as in (16).

(16)   Collaborative writing during this class was mostly very enjoyable, although the effectiveness varied from the person you were working with.

This result corresponds with more students reporting total agreement with the statement "The effectiveness of pair work depends on my partner's personality" at the end of the study: at Time 1, 33.3% ($M = 4.13$; SD = .834); at Time 10, 53.3% ($M = 4.53$; SD = .516). Clearly, a multitude of variables complement or compete with each other in determining the success of collaborative L2 writing at tertiary level. The results reported in this section indicate that intersubjectivity and the attainability of the writing tasks are two crucial determinants.

Collaborative L2 writing and syntactic complexity development

Our final research question asked whether collaborative L2 writing promotes the development of syntactic complexity features characteristic of academic writing. Given the short period of linguistic development captured by the present study, Wilcoxon signed-rank tests revealed significant differences only for two measures of Biber et al.'s (2011) index at stage four.

From Time 1 to Time 10, Group 2 used significantly fewer *adjectival to-complement clauses* ($z = −2.275$, $p = .021$, $r = −.49$). A possible explanation for this result might be that some students adopted a nominalised writing style, as seen in (17), excerpts from Ela's essays at Time 1 and Time 10.

(17)   Time 1:   "since educated people *are likely to be employed* and to receive promising career opportunities"

Time 10:   "Furthermore, education increases *the chance to get employment opportunities*."

It is possible that dyads' negotiations of GM partially account for the statistically significant decrease of adjectival to-complement clauses in Group 2. A complete discourse analysis of transcribed protocols and a qualitative analysis of the essay corpora may reveal more instances supporting this conjecture.

Conversely, in Group 1 the frequency of *prepositional phrases with heads other than 'of' and abstract meaning* significantly increased from Time 1 to Time 10 ($z = −2.429$, $p = .012$, $r = −.52$). This finding supports one of Biber et al.'s predictions of syntactic complexity development in academic writing and corroborates Parkinson and Musgrave's (2014) observation of a higher frequency of this feature in the writing of advanced L2 students. Parallels can also be drawn to the findings of Crossley and McNamara (2014) and Mazgutova and Kormos (2015), both of whom report longitudinal growth of noun-phrase complexity in the academic essays of L2 learners enrolled in dedicated writing courses. In contrast, in the present study students in Group 1 did not receive writing tuition, which might support the contention that extended and unguided academic writing practice and/or task repetition (cf. Nitta & Baba, 2014) may promote genre awareness and the development of lexicogrammatical resources to express abstract concepts.

Conclusion and implications

The study reported in this chapter looked into potential language learning affordances of collaborative L2 writing in digital environments. We focused on the meaning-making dimensions of LREs using an innovative taxonomy comprising the categories of *lexicogrammar problem* and *lexicogrammar alternative*. As such, the study provides novel empirical insights into collaborative writing by uncovering complex meaning negotiations and deep problem-solving behaviour in L2 university dyads. Observed instances of pairs negotiating intricate linguistic structures, such as GM, and engaging in concurrent metalinguistic and ideational thinking suggest a strong connection between collaborative writing and language learning.

In addition, findings from the surveys indicate resoundingly positive student responses, which further endorses the pedagogical value of collaborative L2 writing as a site for language development.

This study does not come without its limitations. One shortcoming is that introspection data (e.g. think-aloud protocols) from individual student writers were not obtained for a comparison of the quality of writing processes and meaning-making behaviours in both writing conditions. This would have allowed us to conclude whether or not collaborative writing at L2 tertiary level engenders a stronger focus on form-meaning relations and deeper levels of language processing with potential learning effects (cf. López-Serrano et al., Chapter 10, this volume). In addition, the study's sole focus on syntactic complexity was at the expense of other important performance parameters (e.g. lexical diversity, accuracy, fluency), their interrelationship, and dynamic emergence. Future research in the domain may approach collaborative writing from a Dynamic Systems perspective to account for the operational interdependence of different performance parameters (see Larsen-Freeman, 2006) and trace *fixed* dyads' "gradually changing meaning-making behaviors and patterns in written language use in terms of their complex heterarchical interrelationships between meaning and form" (Byrnes, Chapter 4, this volume, p. 81).

Future research should be situated in a dedicated writing course populated by first-year university students with little academic writing experience and employ writing tasks that better correspond to participants' educational needs (e.g. abstracts, term paper components). Students may then approach collaborative writing with higher intrinsic motivation, which, in turn, may positively influence language learning outcomes. In genre-oriented composition classes, for example, the teacher-guided joint construction phase of the teaching-learning cycle may be extended to include a collaborative writing component. This would allow students to support each other in choosing appropriate linguistic resources for textual meaning-making, thereby nurturing the transition to competence and autonomy in L2 academic writing.

References

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511920776

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35. https://doi.org/10.5054/tq.2011.244483

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education*, 20, 50–66. https://doi.org/10.1016/j.linged.2009.01.005

Byrnes, H. (2014). Theorizing language development at the intersection of 'task' and L2 writing: Reconsidering complexity. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 79–103). Amsterdam: John Benjamins.

Byrnes, H., & Manchón, R. M. (2014). Task-based language learning: Insights from and for L2 writing: An introduction. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 1–23). Amsterdam: John Benjamins.

Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79. https://doi.org/10.1016/j.jslw.2014.09.006

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7(4), 482–511. https://doi.org/10.1177/0741088390007004003

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.

García Mayo, M. P., & Azkarai, A. (2016). EFL task-based interaction: Does task modality impact on language-related episodes? In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 241–266). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.45.10gar

Göknil, A. (2018). L2 collaborative writing at tertiary level: An analysis of topic sentence formation in academic essays (Unpublished master's thesis). University of Münster.

Halliday, M. A. K. (2002). Spoken and written modes of meaning. In J. J. Webster (Ed.), *On grammar* (pp. 323–351). London: Continuum.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.). London: Edward Arnold.

Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *Modern Language Journal*, 92, 114–130. https://doi.org/10.1111/j.1540-4781.2008.00690.x

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590–619. https://doi.org/10.1093/applin/aml029

Liardét, C. L. (2013). An exploration of Chinese EFL learner's deployment of grammatical metaphor: Learning to make academically valued meanings. *Journal of Second Language Writing*, 22, 161–178. https://doi.org/10.1016/j.jslw.2013.03.008

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York, NY: Academic Press.

Manchón, R. M., & Roca de Larios, J. (2007). Writing-to-learn in instructed language learning contexts. In E. Alcón Soler & M. P. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 101–121). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-5639-0_6

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Martin, J. R. (1989). *Factual writing: Exploring and challenging social reality*. Oxford: OUP.

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15. https://doi.org/10.1016/j.jslw.2015.06.004

McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11(2), 113–128. https://doi.org/10.1093/applin/11.2.113

Nitta, R., & Baba, K. (2014). Task repetition and L2 writing development: A longitudinal study from a dynamic systems perspective. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 107–136). Amsterdam: John Benjamins.

Norrby, C., & Håkansson, G. (2007). The interaction of complexity and grammatical processability: The case of Swedish as a foreign language. *International Review of Applied Linguistics*, 45, 45–68. https://doi.org/10.1515/IRAL.2007.002

Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing*, 29, 82–94. https://doi.org/10.1016/j.jslw.2015.06.008

Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes*, 14, 48–59. https://doi.org/10.1016/j.jeap.2013.12.001

Reinders, H. (2009). Learner uptake and acquisition in three grammar-oriented production activities. *Language Teaching Research*, 13(2), 201–222. https://doi.org/10.1177/1362168809103449

Roca de Larios, J., Murphy, L., & Manchón, R. M. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a foreign language. *Journal of Second Language Writing*, 8(1), 13–44. https://doi.org/10.1016/S1060-3743(99)80111-8

Ryshina-Pankova, M. (2015). A meaning-based approach to the study of complexity in L2 writing: The case of grammatical metaphor. *Journal of Second Language Writing*, 29, 51–63. https://doi.org/10.1016/j.jslw.2015.06.005

Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20, 286–305. https://doi.org/10.1016/j.jslw.2011.05.010

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System*, 27(3), 363–374. https://doi.org/10.1016/S0346-251X(99)00031-7

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158. https://doi.org/10.1111/1467-9922.00179

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14, 153–173. https://doi.org/10.1016/j.jslw.2005.05.002

Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness*, 17(2), 95–114. https://doi.org/10.1080/09658410802146644

Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). Clevedon, UK: Multilingual Matters.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: OUP.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford: OUP.

Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London: Continuum.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371–391. https://doi.org/10.1093/applin/16.3.371

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82(3), 320–337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445–466. https://doi.org/10.1177/0265532209104670

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 312–331. https://doi.org/10.1016/j.jslw.2012.09.007

Appendix A. Essay topics and writing prompts

Time	Topic	Prompt
		Write an academic essay…
1/10	Education	to discuss the significance of education on life quality.
2	L2 Writing Pedagogies	to discuss why teachers should draw on the best of what the different approaches to L2 writing instruction offer.
3	Cognitive SLA	to discuss how psycholinguistic perspectives on second language learning provide a rationale for the use of collaborative writing activities.
4	Sociocultural Theory	to discuss the different dimensions of scaffolding in L2 composition classrooms.
5	Collaborative Dialogue	to explain the different factors that affect the volume and quality of LREs.
6	Patterns of Dyadic Interaction	to discuss the following blog entry by an ESL teacher: "I think it is useful to change the pairing … of students for every activity in the ESL classroom. In general, I find that students who always work with the same partner become lazy and apathetic. Students get bored with the same routine."
7	Learner Perceptions	to explain why some L2 learners may feel reluctant to engage in collaborative writing activities.
8	CW in CMC	to discuss the following statement: "CMC is not necessarily a superior environment for interaction and attention to language than face-to-face interaction" (Storch, 2013).
9	CW Learning Outcomes	to discuss the potential benefits of collaborative writing over solitary writing.

Appendix B. Main survey (items 30–50)

Please tick the boxes which best reflect the extent to which you agree or disagree with the following statements. <u>Only tick one box for each item</u>.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
	1	2	3	4	5
I think writing academic essays in English with a partner would…					
30. … make the writing process easier.	☐	☐	☐	☐	☐
31. … make the writing process more enjoyable.	☐	☐	☐	☐	☐
… motivate me to improve…					
32. … my academic writing skills.	☐	☐	☐	☐	☐
33. … my general writing skills.	☐	☐	☐	☐	☐
…allow me…					
34. … to exchange ideas.	☐	☐	☐	☐	☐
35. … to get useful suggestions for improving my writing.	☐	☐	☐	☐	☐
…help me to…					
36. … learn from my own and my partner's mistakes.	☐	☐	☐	☐	☐
37. … perform better as a student.	☐	☐	☐	☐	☐
38. … make my writing more coherent.	☐	☐	☐	☐	☐
39. … make my writing more cohesive.	☐	☐	☐	☐	☐
… help me to expand my knowledge of…					
40. …academic vocabulary.	☐	☐	☐	☐	☐
41. …general vocabulary.	☐	☐	☐	☐	☐
…help me to improve…					
42. …the assessment of my writing.	☐	☐	☐	☐	☐
43. …my general grammar knowledge.	☐	☐	☐	☐	☐
44. …the sentence structure of my writing.	☐	☐	☐	☐	☐
45. …the accuracy of my writing.	☐	☐	☐	☐	☐
46. …my spelling.	☐	☐	☐	☐	☐
47. …the structure of my essays.	☐	☐	☐	☐	☐
48. …the content of my essays.	☐	☐	☐	☐	☐
49. …my revision skills.	☐	☐	☐	☐	☐
50. …my editing skills.	☐	☐	☐	☐	☐

Appendix C. Retrospective survey

Name:_____

Date:_____

Please <u>briefly</u> answer the following questions to reflect on today's writing assignment:

1. How did you feel working with your partner?[2] ....
    ....
    ....
    ....
2. What worked well? ....
    ....
    ....
    ....
3. What didn't work so well? ....
    ....
    ....
    ....
4. What did you gain form the activity? ....
    ....
    ....
    ....

---

2. Only for Group 2 (pair work)

Writing on history in a Content and Language Integrated Learning (CLIL) context

Development of grammatical metaphor and abstraction as evidence of language learning

Rachel Whittaker and Anne McCabe

Universidad Autónoma de Madrid / Saint Louis University, Madrid Campus

In this chapter we attempt to answer Manchón and Williams' (2016, p. 580) question as to "how language and writing develop in content-based instruction and CLIL programs" as a key concern in the study of writing as a site for language learning. Our longitudinal study from a CLIL context traces students' development in writing in English by analyzing a key feature of disciplinary literacy, grammatical metaphor (GM), created through nominalizations and abstract nouns. Data for this study consists of 64 texts by the same 16 students, on a topic from the history syllabus, collected yearly over the four years of obligatory secondary education. In the data we identified and analyzed just under 1,500 instances of GM. Results showed a significantly higher use and greater accuracy of GM in second cycle texts. Given the role of writing in learning subject discourses, the chapter closes by reflecting on the implications of these findings for practice and for research directions for writing in CLIL contexts.

Introduction

The context for writing examined in this chapter is that of Content and Language Integrated Learning (CLIL), a growing phenomenon in Europe (European Commission, 2017), involving students in learning one or more content subjects through a foreign language (FL). This policy aims at improving Europeans' competence in the languages of the EU, by increasing contact time and communicative use of the FL. For the content teachers involved, the priority tends to be spoken language for classroom interaction, while writing rarely appears (Dalton-Puffer, 2007) except at examination time. The questions we ask in this chapter are: How do students respond to the challenge of writing the discourse of the discipline? Do they develop, over time, the language features required to express the knowledge the

discipline creates? We are looking, then, for a way to answer Manchón and Williams' (2016, p. 580) question as to "how language and writing develop in content-based instruction and CLIL programs" as a key concern in the study of writing as a site for language learning (See also Introduction to this volume).

Our study is located in history classes taught in English in a Madrid state school. The corpus consists of written data collected as a yearly end-of-topic revision task over the four years of obligatory secondary education (Grades 7–10). We take a quantitative and qualitative approach in analyzing the data, to trace the development of key meaning-making resources for the subject area identified by the Systemic Functional Linguistics (SFL) model: grammatical metaphor and abstraction. Their use allows historians to move from telling historical 'stories', set in chronological time, to creating texts which consider cause-effect relations, and interpret and evaluate history (Coffin, 2006; Eggins, Martin & Wignell, 1993; Lorenzo & Dalton-Puffer, 2016).

The study sees writing as contributing to language learning in the students' increasing ability to use these features, as well as to produce more correct forms over time (see Chapter 4, this volume). The analysis allows us to trace the students' linguistic development not only in the measures of fluency and accuracy, but especially in terms of complexity (see Wolfe-Quintero, Inagaki, & Kim, 1998 for these CAF measures in L2 writing, and Polio & Park, 2016 for a recent overview). In this case, however, we refer to meaning-based complexity (Ryshina-Pankova, 2015, p. 52), in terms of content-specific meanings, that is, abstraction in the representation of history.

In what follows we first discuss studies of writing in secondary CLIL contexts and then turn to the concept of grammatical metaphor and its role in the construction of historical discourse. We then describe the data collection and analysis. Results are considered in relation to foreign language learning seen as the expansion of students' resources for making historical meanings, while they also move towards more accurate forms as they write in English. In the conclusion we propose integrating explicit work on subject-specific writing into the teaching of subjects in a foreign or second language, and consider research directions for the field.

Writing in CLIL classes and academic language development

Several features of writing in an FL give this productive activity potential for language learning. Manchón's (2011) review of research into writing to learn an additional language points to some common characteristics in studies from cognitive or social research frameworks: Learners are involved in communicative tasks requiring writing, during which they notice gaps or difficulties with the language and try to solve these problems. At the same time, the stability of writing allows writers to reflect

on the elements of language visible to them, rehearse them, and manipulate them as they strive to construct meaning (Manchón & Williams, 2016; Williams, 2012).

Additionally, research suggests that the cognitive effort involved in L2 writing in subject areas leads to learning content (Coetzee-Lachmann, 2007; Hirvela, 2011). Teachers find that the activity of creating a written text on subject content helps to "*develop* understanding" (de Oliveira, 2011, p. 41, italics original), and students themselves report that it organizes information and fixes it in the memory.[1] From a psychological perspective, Heine's (2010) study revealed that the effort of writing in a FL leads to deeper processing, which has been interpreted as influencing the learning of both content and language. Thus, "bilingual classrooms are acquisition-rich environments in which learners are necessarily engaged in the manipulation of complex language" (Lorenzo & Rodríguez, 2014, p. 65), conditions that are propitious for learning an FL (cf. de Graff, Koopman, & Westhoff, 2007; Meyer & Coyle, 2017). A subject area such as history, for example, offers opportunities to involve learners in thinking at different levels, such as recalling knowledge, explaining a given phenomenon, or evaluating the impact of a historical character or event (Beacco, 2010; Coffin, 2006; Lorenzo & Dalton-Puffer, 2016). If, as Byrnes (2011, p. 148) argues, "among the key insights learners must gain about a language system is its meaning-making resource quality rather than its rule-based quality", then these cognitive demands represent opportunities to extend students' meaning-making potential (see also Byrnes's Chapter 4, this volume).

Turning to writing and language development, several cross-sectional and longitudinal studies have analysed the development of writing in CLIL contexts. Comparing CLIL and EFL student writing, cross-sectional studies applying various CAF measures and/or using holistic bands of features to evaluate texts (Jexenflicker & Dalton-Puffer, 2010; Lasagabaster, 2008, 2011; Navés, 2011; Ruiz de Zarobe, 2010) have found greater improvement for CLIL pupils one or two grades younger than the EFL groups. As to comparative longitudinal studies, two data collections separated by only one academic year found significant differences between a CLIL plus EFL group and an EFL-only group in accuracy (Roquet & Pérez-Vidal, 2015), but not in other production measures. However, a longer study comparing CLIL and non-CLIL groups over a three-year period found significant improvement on most of the different CAF-type features measured for CLIL students' texts, while the EFL students only improved significantly in accuracy and lexical variety (Gené-Gil, Juan-Garau, & Salazar-Noguera, 2015, p. 287).[2]

---

1. Personal communication from a history teacher and researcher.

2. One cannot help recalling Craig Chaudron's smiling comment that students tend to learn what they are taught.

In studies of writing within a content subject, the emphasis is often on the development of features of academic language, rather than on the CAF features of the texts produced. Qualitative, discourse analysis-based studies (Llinares & Whittaker, 2010; Whittaker & Llinares, 2009, 2011) show that different ideational, interpersonal, and textual features begin to emerge appropriately in texts written by CLIL groups studying history in English in obligatory secondary education, albeit more slowly than in those written by L1 students of the same age. Thus Järvinen's (2010) study of features of academic register in a data collection of Grade 8 Finnish CLIL and international students' history texts found higher lexical density and fewer clauses per sentence in the writing of the international students, who received education entirely in English. Interestingly, a comparison of the CLIL students' texts in English and in Finnish showed parallels in the development of written register: either high or low lexical density in both languages. Even after only one year of learning subjects through an FL, writing on content has been found to have a very positive effect on language development. In this respect, Lorenzo and Moore (2010, p. 33) report on "frequent examples demonstrating clear emergent abilities with regard to more complex grammar" and other advanced features of language in short paragraph answers on content topics, which they interpret as an effect of the cognitive challenges posed by the writing tasks. The creation of textuality also develops through studying content in a FL, as evidenced by results of a quantitative longitudinal study of cohesion and coherence in texts written on history topics over the four years of secondary education (Whittaker, Llinares, & McCabe, 2011). A longitudinal study of the creation of the 'voice' appropriate for a task of writing in history also showed students developing control of the linguistic resources of evaluation, allowing them to present events in a suitably objective way, or to evaluate historical figures for their moral and other characteristics (McCabe & Whittaker, 2017).

Finally, Lorenzo's detailed studies of academic language in FL and L1 history classes (see especially Lorenzo, 2016; Lorenzo & Rodríguez, 2014) provide empirical evidence of the development of features of student writing in CLIL secondary history classes. These researchers found that, although the area of lexical richness showed most dramatic gains as the students moved through Grades 9 to 12, the grammatical patterns in the texts also changed significantly. Among the structures which developed, modification inside the nominal group stands out for its steady increase, an increase that was statistically significant in the final year in relation to previous years. It is to this category that we now turn, examining its functional role in academic language as a manifestation of grammatical metaphor.

The role of grammatical metaphor (GM) in meaning-making in history writing

Grammatical metaphor (GM) is an important meaning-making resource, as it "leads to an *expansion of the meaning potential*: by creating new patterns of structural realization, it opens up new systemic domains of meaning" (Halliday & Matthiessen, 2014, p. 699; emphasis original). It has been studied extensively by researchers using the SFL framework given its central role in academic language, and so in educational success. The linguistic resources available to language producers through GM allow for an increased variety in functional expression, and so GM is linked to development of further meaning-making abilities in communication.

While GM resources exist across the three meta-functions of language, we mainly focus on ideational GM for its role in transforming the congruent grammar of speech into forms which allow the building of abstract knowledge (Byrnes, 2006; Christie, 2002, 2012). Ideational GM draws on the system of transitivity (processes, participants, and circumstances), on the logical connections between clauses and sentences, and on elements at the level of group (especially nominals); it operates between grammatical categories, rather than between lexical items as in lexical metaphor. Figure 1 shows congruent and metaphorical versions of similar meanings in a set of invented examples:

Congruent version	Metaphorical version
During WWI, soldiers used more weapons when they fought, so they destroyed many things.	*WWI saw an increase in technology for warfare, which led to great destruction.*
– participants are expressed as nouns (*soldiers, weapons, things*) and refer to concrete phenomena	– concrete participants disappear in favour of abstractions (*technology*)
– processes are expressed as verbs (*used, fought, destroyed*)	– processes are expressed as nouns (*warfare, destruction*)
– temporal location (a circumstance) is expressed through a prepositional phrase of time (*during WWI*) or through a temporal subordinating conjunction (*when*)	– temporal location is given a name, as a noun (*WWI*), which becomes a participant, as subject of the clause
– comparison of quantity is expressed through a quantifier (*more*)	– comparison of quantity is expressed through a noun (*increase*)
– cause/consequence is expressed through a conjunction (*so*)	– cause/consequence is expressed as a verb (*led to*)

Figure 1. Congruent and metaphorical realizations

On the left side of Figure 1, there is a match, or congruence, between the meanings and their expression in language categories; for example, nouns express participants in the clause. On the right-hand side, however, that congruence between meaning and linguistic categories is broken, for example through the expression of processes as nouns. This move between categories is called grammatical metaphor, since it involves "a semantic junction or combination of two meanings" (Ryshina-Pankova & Byrnes, 2013, p. 187), those of the original and of the new grammatical category.

While different classes of word can play a part in GM, the noun is the main protagonist in the ideational domain. As Christie (2012, p. 110) has found in studies of language development in schooling, GM "enables the student of history to turn activities into phenomena – actually grammatical participants – about which subsequent explanation and interpretation can be developed". Indeed, a motivating factor for such reconstruals is that they can take on the meaning-expansion resources applicable to nouns: "when 'somebody remembering something' is reconstrued as 'memory' it can be classified and characterised just like other entities" (Halliday & Matthiessen, 2014, p. 714). Also, as GM can turn happenings into abstract phenomena and events, historians are "… able to take language out of its immediate context – i.e. 'abstract' or 'distance' language from the then-and-there" (Eggins, Martin, & Wignell, 1993, p. 96), and so discuss, interpret and evaluate these now abstract representations. Evidently, GM is not a language resource which pupils have readily at their disposal; rather, it is a linguistic and rhetorical ability which needs developing.

The development of grammatical metaphor

In mother-tongue language learning contexts, ideational GM is "associated with the discourses of education and science, bureaucracy and law" (Halliday & Matthiessen, 2014, p. 709) and so requires access to formal contexts to develop. Researchers have suggested a development path for GM acquisition throughout the life of young children and through the years of schooling. In English mother-tongue contexts, Halliday (1993) considers that children begin contact with ideational GMs around the age of nine, which Christie (2012) corroborates, although her research shows that variability is the norm. Given the demands of secondary school, by at least mid-adolescence, students need to demonstrate control of GM (both decoding and encoding) as part of a successful move into disciplinary discourses. In Halliday's studies of the ontogenesis of language, control of GM is one of the stages of language learning:

> As grammatical generalization is the key for entering into language, and to systematic commonsense knowledge, and grammatical abstractness is the key for entering into literacy, and to primary educational knowledge, so grammatical metaphor is the key for entering into the next level, that of secondary education, and of knowledge that is discipline-based and technical.   (Halliday, 1993, p. 111)

It could be argued, then, that any account of development of student writing in a CLIL program in secondary education needs to trace the development of this key resource.

In non-L1 language learning contexts, at university level, studies in a genre-based FL programme led by Heidi Byrnes at Georgetown University (Byrnes, 2009; Ryshina-Pankova, 2015; Ryshina-Pankova & Byrnes, 2013) show development of GM by students learning German, following explicit instruction; Colombi (2009) demonstrates similar results for Spanish heritage learners. Liardét's (2013) cross-sectional study of first and fourth year L2 Chinese university students learning English, with no explicit instruction of GM, found more successful use in later essays, albeit with "unrealized potential" (2013, p. 176).

In the absence of explicit instruction, it seems that development of GM must depend on repeated exposure to more written and formal varieties of texts (Christie, 2012; Christie & Derewianka, 2008; Halliday, 1993; Liardét, 2015; Schleppegrell, 2004). As we have seen, studies on the development of writing in secondary CLIL contexts point to the increasing use of pre- and post-modified nouns, which provides indirect evidence for the development of GM (Järvinen, 2010; Lorenzo & Dalton-Puffer, 2016; Lorenzo & Rodríguez, 2014; Whittaker et al., 2011). Our longitudinal study of student writing provides direct evidence for the development of GM over the years of secondary schooling in a CLIL history context.

The study

Research questions

Our overall research question for the CLIL secondary school students writing in history was whether their texts showed development in use of grammatical metaphor (GM). This was broken down into more specific questions:

1. Do the texts increase in lexical density?
2. Do the students use GM (nominalization or abstraction)?
3. Do they use more of these resources in later years?
4. What types of resources do they use?

5. Do they use more types in later years?
6. Do the formal features in using these resources show differences in later work by the students?
7. Do they show differences across texts rated differently?

Method

Context, participants, and data collection procedures

The study, part of a larger project on language needs in school subjects taught through the medium of English (see UAM-CLIL research group at uam-clil.org), was carried out in a state secondary school in Madrid. Participants were in the bilingual section of the school, with at least two content subjects taught in English during the four years of obligatory secondary education. The project took an exploratory approach, analysing a corpus of classroom data including whole-class oral interaction, a sample of individual student's oral production, and unsupported in-class writing, with the aim of evaluating the extent to which the pupils were able to take part in meaning-making activities leading to learning the content of the subject. Thus there was no intervention at any point. Classes focused on the content of the very wide history syllabus, with work on language limited to technical lexis for the subject, such as names of periods or definitions of concepts. For the data collection sessions, prompts were designed by the research group after analysing the history syllabus and in consultation with the subject specialists. Though writing was not a part of the normal activity in the teaching/learning cycle, the teachers enthusiastically agreed to its inclusion.

The once-a-year data collection sessions consisted of recording a whole-class end-of-topic summary session, which was followed some days later by an individual writing task under test conditions covering the same material (Grade 7 on ancient civilizations, Grade 8 on feudal times, Grade 9 on Philip II, and Grade 10 on World War I), for which 20 minutes was allowed. To trace the longitudinal development of GM, we analysed texts by the same 16 students over the four years.

In order to assess the use of GM as a marker of quality in writing, as one feature of success of a written text, an expert CLIL history teacher was asked to rate the essays holistically, by classifying them into three categories: high, average, and low (cf. de Oliveira, 2011). The rater based her classification on expected achievement for the class in that grade. Table 1 provides a general description of the corpus studied, in word counts and in the rating of texts for each cycle.

Table 1. Description of the corpus

	Cycle 1 Grades 7 & 8	Cycle 2 Grades 9 & 10	Totals
Number of texts	32	32	64
Number of words	4782	6263	11045
Lower-rated texts	13	7	20
Average-rated texts	9	7	16
Higher-rated texts	10	18	28

Data analysis procedures

Following previous studies of grammatical metaphor (Liardét, 2013; Ravelli, 2003; Ryshina-Pankova & Byrnes, 2013), we analysed the data for two classes of GM: nominalization and abstract nouns. Nominalization refers to processes, attributes, or circumstances which are reconstrued as things or events. Two principles guided identification of nominalizations: derivation and agnation. In derivation, the incongruent form is recognized through a productive suffix, for example, 'obligation' – 'to obligate oneself'; 'importance' – 'important'. Agnation refers to a metaphorical form which has a corresponding congruent form, through conversion, as in 'to attempt' – 'an attempt'. Abstract nouns, on the other hand, do not reveal a trace to the original process or quality. For their identification, following Schmid (2000), we considered the context: These nouns were functioning as "abstract 'containers' or 'shells' for some kind of content" (Ryshina-Pankova & Byrnes, 2013, p. 187); examples from our data are 'problem', 'reason', 'consequence' or 'fact'. In some studies of GM, nominalization forms the sole focus of analysis; however, there are several arguments for also including abstract nouns. First, not all nominalizations can be thought of as processes encoded through nouns; that is, not all encode a configuration of events which can be unpacked to show elided participants taking part in a process. In Fontaine's (2017) examples "The examination of the patients took a long time" and "The examination was on the table"; the latter use of 'examination' does not encode covert participants in the same way as the nominalization 'examination' in the former instance does, where it implies a doctor as agent. Ryshina-Pankova (2015, p. 54) also points out that "because GM is first and foremost a semantic concept that refers to the incongruity in realization, it cannot be equaled with nominalizations".

Second, the notion of incongruence as the deciding factor for GM leads to a consideration of other "metaphorical abstractions", or "ready-made abstract lexical items that summarise a series of happenings or an entire situation, such as the term *war*" (Painter, Derewianka, & Torr 2005, p. 584). Therefore, a "fuzzy boundary" exists between abstract nouns and 'pure' nominalizations (as GM), which is

not surprising "as one is a step towards the other (as confirmed ontogenetically)" (Ravelli 2003, p. 60). Thus, any study of GM must consider its purposes.

In light of this fuzzy boundary, there are two factors which led us to include abstract nouns as instances of GM in our study: The first is our interest in development – here, whether students' texts grew in increasing abstraction over the four years; the second is related to the technicality of history, which is constructed through both nominalizations and abstract nouns, a point which we return to in the discussion of the results. At the same time, in the examples from Figure 1, the move from the more concrete 'weapons when they fought' to the more abstract 'technology for warfare' helps to make the point that technical abstract nouns provide the same kinds of distancing from lived experience (Ryshina-Pankova, 2015), with their packaging up of whole events, as do nominalizations. And it is through this distancing that the voice of the historian emerges. The two types of GM studied are listed in Table 2, with examples from the data.

Table 2. Categories of grammatical metaphor

Category (Coding tag)	Explanation	Examples
Nominalization	Process behind nominalized form is recoverable	*mercantilism, evolution*
Abstract noun	Process is not recoverable from the word form, which refers to an abstraction	*agriculture; society*

Given our interest in language learning through CLIL writing, formal features were also analysed. In the case of errors, instances were classified for location and type of error. For this we adapted Liardét's (2013, 2016) system, and divided errors into two types: head-noun errors, occurring in the head noun itself, and "co-text intermediacy" errors, occurring elsewhere in the same nominal group. Examples (1)–(5) show head-noun errors from our data[3]:

(1) gave them the correct <u>earn</u>   <G7-S14>

(2) a very good <u>work</u>   <G8-S9>

(3) pay for the <u>damages</u>   <G9-S13>

(4) explosion of new <u>invents</u>   <G10-S7>

(5) her <u>increasing</u> of power   <G10-S17>

---

3. All examples from the data show the grade (G) and then the student number (S), and are represented verbatim for spelling, punctuation, etc.

Examples (6) and (7) are instances of co-text intermediacy errors:

(6)   the ancients civilitations   <G7-S17>

(7)   obligation of worked the lands   <G8-S7>

In Example (6), the head noun (while misspelled) is a correct form, yet the student pluralizes the classifier 'ancients'. In (7), again, the head noun 'obligation' is correct, yet the student writes 'worked' instead of using the gerund form after the preposition.

For the study, a network with the categories of GM and error types described above was created in the UAM CorpusTool (O'Donnell, 2008). Two raters (the authors) independently tagged each instance of GM identified in the data, and then resolved any doubts or disagreements through discussion (see Byrnes, 2009). Coding was recorded using CorpusTool. Considering results of other longitudinal studies of CLIL writing (Gené-Gil et al., 2015; Roquet & Pérez-Vidal, 2016), and given the variability in language development, the results were calculated using a two-year cycle, Cycle 1 encompassing Grades 7 and 8 of secondary schooling and Cycle 2 Grades 9 and 10. For calculations, CorpusTool's statistical package was used. Given the role of lexical variety in CLIL writing development (Lorenzo & Rodríguez, 2014), raw counts of the different lexical types are included.

Results

We begin with lexical density (as a percentage of lexical words to total words) as a more general measure of the data, given its relationship to GM in creating a more written academic register (Christie, 2012). Table 3 displays the results across cycles and across quality ratings of the texts:

Table 3. Lexical density in cycle 1 and cycle 2 by text quality

	Cycle 1	Cycle 2
Lower-rated	47%	47%
Average-rated	45%	48%
Higher-rated	46%	51%

While the lower-rated essays show no difference across cycles in lexical density, the average-rated texts increase by 3% and the higher-rated by 5%.

Comparison of the use of GM in the two cycles appears in Table 4. Both raw numbers and per 1,000 tokens[4] are given with the t value and significance.

Table 4. Grammatical metaphor use: Cycle 1 vs. Cycle 2

	Cycle 1		Cycle 2		TStat	Signif.
	N	per 1000 tokens	N	per 1000 tokens		
Nominalization	184	33.58	255	36.34	6.56	$p < .02.$
Abstract noun	234	42.70	723	103.02	6.62	$p < .02.$

Table 4 shows a significant increase of GM through both nominalizations and abstract nouns in the second cycle.

Turning to variety in the GM types, given the variability in text length and different topics covered in the study, raw counts of the lexical types found in the texts over the four years are shown.
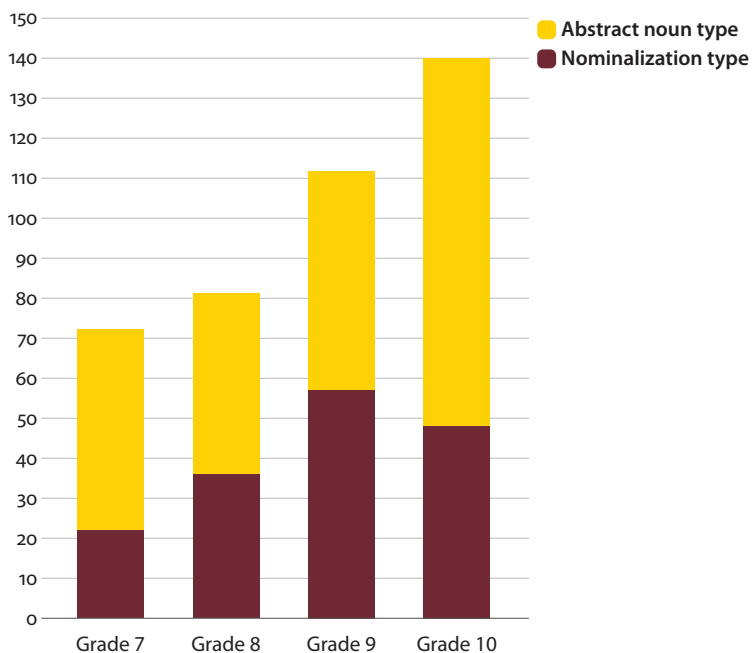
Figure 2. Variety of grammatical metaphor types Grades 7 to 10

---

4.   UAM Corpustool calculates results per number of tokens, which includes words and all punctuation.

Figure 2 shows a steady increase in nominalization types up to Grade 9. The number remains equal into Grade 10, where there is a marked increase in abstract noun types. Looking in detail, repetition and use of lexis from the prompts decreases over the four years. While in Grade 7 the nominalization type most often repeated (46 out of 184 tokens) is a variation of the word 'civilizations', other GM items ('government', 'development', 'death', 'power', 'society') often appear. In Grade 8, repetition decreases considerably, and examples of GM are 'population' 'work', 'remuneration', 'agriculture' and 'consequences'. In Grade 9, the number of types continues to expand, and students start using places with an abstract meaning rather than as location, a feature of history writing, as in Example (8)

(8)   With mercantilism, Spain became self-sufficient…   <G9-S21>

This trend continues into Grade 10, when the writers very frequently encode locations, such as Germany, Austria, France and Britain, as participants. In this year, frequent nominalizations include 'assassination', 'blockade', and 'alliances', and abstract nouns include 'treaty', 'war' and 'reasons'.

Table 5 presents the results for accuracy across the two cycles.

Table 5. Grammatical metaphor form: Cycle 1 and Cycle 2

	Cycle 1		Cycle 2		TStat	Signif.
	N	per 1000 tokens	N	per 1000 tokens		
Correct	347	63.32	934	133.09	6.99	$p < .02.$
Error	89	16.24	79	11.26	6.99	$p < .02.$
Co-text-intermediacy	58	10.58	52	7.41	5.44	$p < .02.$
Noun-form	30	5.47	27	3.85	3.80	$p < .02.$

In Cycle 2, there is a significant increase in the correct forms of the GMs, and a correspondingly significant decrease in errors of both types, in head noun forms and elsewhere in their modifiers. Errors are more frequent in the modifiers than in the head noun.

A possible relationship between GM and perceived quality of the texts was examined by comparing the lower- and higher-rated texts. Tables 6 and 7 display the comparison of lower- and higher-rated texts for both GM type and for accuracy in the two cycles.

Table 6. Grammatical metaphor in cycle 1: Lower- vs. higher-rated texts

GM type	Cycle 1 Lower		Cycle 1 Higher		Tstat	Signif.
	N	per 1000 tokens	N	per 1000 tokens		
Nominalization	77	44.66	71	32.38	1.506	
Abstract noun Form	73	42.34	98	44.69	1.789	$p < .1.$
Correct	117	67.87	148	67.49	2.409	$p < .02.$
Error	40	23.20	26	11.86	2.409	$p < .02.$

Table 7. Grammatical metaphor in cycle 2: Lower- vs. higher-rated texts

GM type	Cycle 2: Lower		Cycle 2 Higher		Tstat	Signif.
	N	per 1000 tokens	N	per 1000 tokens		
Nominalization	33	32.13	167	36.35	2.447	$p < .02.$
Abstract noun Form	56	54.43	529	115.15	3.016	$p < .02.$
Correct	79	76.92	670	145.84	3.476	$p < .02.$
Error	16	15.58	48	10.45	3.476	$p < .02.$

In Cycle 1, there is little difference in GM between the lower- and higher-rated texts, apart from a trend towards more abstract nouns in the higher group, which is also significantly more target-like in the forms of GM. In Cycle 2, however, both classes of GM are significantly more frequent in the higher-rated group, which again shows a significantly greater use of correct forms. Furthermore, the use of place as abstraction, an interesting feature of history discourse, was significantly higher in Cycle 2 (finer-grained analyses not reflected in the table for space reasons).

Summing up these results, and in answer to our research questions, the average- and higher-rated texts increase slightly in lexical density. As to GM use, from the first to the second cycle, nominalization and abstraction increase significantly. In terms of lexical variety, nominalization types show a steady increase while abstraction types grow yearly, especially in the final year. Analysis of correct use of the features studied show significant increases in accuracy, both in head nouns and the whole nominal group, in the second cycle. Relations between quality and GM in the first cycle point to the role of errors rather than GM use in the rating, while in the second cycle both of the GM types are significantly more frequent in higher-rated texts, which also have a significantly lower count in errors.

Discussion: Expanding meaning-making resources through writing on history

In an attempt to shed light on the connection between writing and language learning, this study set out to investigate grammatical metaphor as an indication of writing development. The increases in number, variety, and correctness of form in lexical density, nominalization, and abstraction that occurred over the four years of data collection point to the important role of writing in bringing about language development, as these meaning-making resources differ from those of speech (see Byrnes, Chapter 4, this volume), and can only develop through interaction with written texts, both through their consumption and, especially, through their production.

First, as regards lexical density, both the average- and higher-rated texts increased in this measure from first to second cycle, while that of the lower-rated texts showed no change. Lexical density is an indicator of more advanced, informationally dense writing, and similar results have been found in studies of writing in history in CLIL contexts (Järvinen, 2010; Lorenzo & Rodríguez, 2014).

Related to density, and possibly overlapping with it, the significant increase found here in both categories of GM, nominalization and abstract nouns, in the second cycle texts has been seen in other studies of FL writing. In fact, Lorenzo and Rodríguez (2014, p. 68) relate the increase in "complex nominals per clause" in higher grades to the development of GM. Their study measuring the trajectory of a large number of features found this category especially significant in Grades 11 and 12.

At the same time as tokens of GM increased significantly in the second cycle texts, lexical resources expanded and became significantly more target-like. In these aspects, some interesting comparisons can be made with Liardét's (2013, 2015, 2016) research on the development of nominalization in Chinese university students writing in English. In our study, by Grade 9, Spanish students had doubled the number of types they produced to create meaning in history. For the Chinese students, too, the activity of writing on academic topics produced a considerable increase – almost threefold – in nominalization types over the four semesters of first and second year (Liardét, 2016, p. 21).

In our attempt to capture development in FL writing, we examined 'intermediate' (Liardét, 2015) stages of GM, instead of discounting non-target forms. Results show how the apprentice writers are experimenting as they build meaningful text, pushing their interlanguage towards the language of history by expanding their nominal groups, as the location of errors shows. While a third of the errors were related to the form of the head noun, it was the modifiers (co-text-intermediacy) which concentrated more non-target forms. That the students' language was developing is clear from the significant drop in errors in the second cycle texts.

Overall, our study corroborates the very important role of nominalization in development in school writing, as highlighted by Lorenzo and Rodriguez (2014, p. 68): "nominalization has been seen as a borderline that marks linguistic adulthood, a feature of expository language and scientific discourse, and this structure seems to reach its full extent in late adolescence". Nominalization and abstraction are the basis of the "technicality of history" (Eggins et al., 1993), and play an important role in its discourse, as our writers are beginning to discover.

As there was no instruction in relation to GM, or to writing in general, we speculate that studying history in an FL provided strong input for writing development, with several characteristics of the discourse of history calling for its use. First, it allows for nominalising actions, as Example (9) shows:

(9)   … It is diferent to our days society, because there the king was the person who has all the power and <u>control</u> all the persons, in our days, all the persons have the same rights and <u>the control</u> are in hands of a lot of people not only in one person.   <G7-S16>

This first-year writer first uses 'control' as a process, and then introduces the nominal form as an abstraction. Variability, typical in language learning, can be seen in students' development towards GM, as shown in examples (10)–(14) from Grade 8. Examples (10) and (11) show more everyday construals of events, with nouns encoding participants and verbs encoding doings:

(10)   they didn't clean their bodie all the days   <G8-S13>

(11)   they didn't have to pay the lord for the lands   <G8-S20>

while Examples (12)–(14) show students moving towards the discourse of history through the use of the abstract noun 'hygiene', as a classifier in (12) and as a head noun in (13) and (14), to encode the activity of people washing, now institutionalized, as a generalization about society:

(12)   …thanks to the very bad health and hygen conditions   <G8-S21>

(13)   The causes the plague was because of the hygene   <G8-S19>

(14)   The hygien wasn't be very good.   <G8-S16>

In (12)–(14), the students show awareness of the abstract noun 'hygiene', albeit with instability in control over its form. Such instability should not be viewed from a deficit perspective, as "often, learners' early attempts at non-congruent forms such as nominalization are clumsy and awkward" (Mohan & Beckett, 2001, p. 426). For development, learners need to experiment, using less congruent, more abstract forms, even when that means instability in the control of those forms. In studying

history through English, the students have been exposed to its nominalizations and technical abstractions, which they begin to incorporate into their own texts.

Furthermore, nominalizations in history writing allow 'doings' to act and to be acted upon:

(15) There was <u>a price revolution</u> too. and it led to build banks.   <G9-S17>

(16) USA with his president Wilson started to made fourteen points to improve <u>Europe relations</u> for future.   <G10-S10>

In examples (15) and (16), the 'price revolution' takes an active role in the building of banks, and 'Europe relations' becomes object of an action. In the corpus, however, most nominalizations appear in relational clauses, where they are given attributes rather than being active participants in the clause; thus, these writers still need to develop this ability more fully.

History writing also provides the specific sites of time, location and cause, presented as abstractions which can be given different grammatical roles, making them incongruent participants in the clause. History obviously revolves around setting in time, and students learn to label historical periods or phases (Coffin, 2006; Schleppegrell, 2011). In Example (17), a Grade 9 writer uses segmented time (underlined) as part of the argument to support an evaluation of Phillip II:

(17) he had <u>3 phases in his evolution</u>, first he had <u>a bad period</u> but then he started to have <u>a time of explendour</u>, he won the battle of Lepanto and he prospered in economy. the society was divided in different groups, So that is not a good point but I think he was a good monarch, he becames the king of Portugal. Then there was <u>a decline</u>.   <G9-S13>

Using relational clauses, the writer divides Phillip II's reign into good and bad periods, then refers to a process related to Phillip II (he became the king) and finally returns to segmented time with an existential clause to introduce a period of decline. Coffin (2006) suggests that students who are further along the developmental pathway in writing history use fewer resources for construing time as setting (e.g. 'In 5000 BC') and more resources for segmenting time, moving towards a more institutionalised understanding of time, which is more abstract and distant from lived experience.

A second site for abstraction is place, which, like time, can function in the clause more congruently as a circumstance, such as 'in France'. However, through reference to geo-political roles of countries, place becomes removed from everyday experience (Veel & Coffin, 1996, pp. 214–215). Thus, place in history loses its sense of concrete physical location and becomes an abstraction, which then can function as metaphorical subject of different semantic classes of verb: action verbs

(Examples (18) and (19)), and verbs of feeling (Example (20)). This use of countries as presenting human feelings is a feature of history writing (Coffin, 2006; McCabe & Whittaker, 2017).

(18)   Germany attacked Belgium   <G10-S1>

(19)   England built a large navy sourrounding Germany to block them <G10-S14>

(20)   when Germany wanted to expand   <G10-S13>

A veritable explosion of this use of place is found in Grade 10 texts, when students wrote about the countries involved in World War I using personification. Example (21) shows a build-up of GM, with the quality 'anxious' turned into the nominalized form 'anxiety' attributed through post-modification to Germany, which then functions as Actor of a process in a non-finite clause:

(21)   the mos important cause was the anxiety of Germany to build up a great Empire and control all Europe with his army and navy.   <G10-S21>

Example (21) shows another semantic area, the logical relation of cause, which plays a key role in history texts. Nominalized 'cause' and the abstract noun 'consequence' make their appearance mainly in Grade 9. In Grade 10, some developments in relation to the expression of cause are noted. First of all, there are more expression types: In addition to 'cause' and 'consequence', students also write about 'reasons' and 'factors'. Secondly, in Grade 10, and not before, students bring evaluation into the nominal group, with Appraisal resources of Attitude: Appreciation (e.g. 'important') (Martin & White, 2005) and Graduation (e.g. 'most'), as in 'the most important cause'.

Of course, simply increasing the use of GM does not always lead to improved quality of a text (Liardét, 2016; Ryshina-Pankova & Byrnes, 2013). An increase in density can lead to a decrease in clarity, as all readers have experienced. Evidence of real growth in ability would be found in texts which also rephrase examples of GM into more everyday language, creating semantic waves of abstraction and concreteness (Maton, 2013; Schleppegrell, 2011). While this was not a focus of the analysis, examples (22) and (23) show this type of rephrasing, indicating understanding of the GM (underlined) used, as well as the flexibility to choose appropriate realizations in the foreign language.

(22)   In 1917 Germany took the <u>decission</u> to start a submarine <u>warfare</u> that means to attack everything in the sea neutral ships, etc.   <G10-S10>

(23)   The war was so long that anyone was available to won the war, this was called <u>deadlock</u>.   <G10-S25>

Clearly, writing on history in English provided many opportunities for students to expand their linguistic resources, as seen in the development of control of GM and abstraction to make subject-appropriate meanings in their texts.

Conclusions and pedagogical implications for writing in CLIL classes

The quantitative and qualitative analysis of our longitudinal corpus provides a unique 'window' onto the role of writing for foreign language development in a European CLIL context, showing how the students' attempts to make meaning set in motion the build-up of grammatical metaphor and so the start of a move towards the lexically dense and abstract register of the subject, history. In the corpus, over the four years we have seen students' texts begin to develop the nominalized grammatical organization of disciplinary writing with the use of GM, as well as to increase in accuracy and in lexical variety related to subject content. It would seem then, that the advantages shown for writing to learn language, permanence and time for reflection (Manchón & Williams, 2016), can well be exploited in contexts in which students learn content through an additional language. These contexts are a site for cognitively demanding writing tasks, in which students can engage as part of the learning of a subject. We can expect, then, that incorporating writing systematically into these classes would lead students to discover gaps in their language knowledge, and explore and expand their resources, with this process of reflection and rewriting leading to "deeper learning", as reported by participants in Manchón and Roca de Larios' (2011) study.

However, to incorporate writing as an integrated part of learning both subject and language, teachers need access to the resources required to create text in a specific educational context, as advocated by Byrnes (2011; this volume). Projects like the Council of Europe's "Languages in Education, Languages for Education" encourage approaches to content learning based on making explicit the language for subject literacy and "the rhetoric of knowledge" (Beacco, 2017, p. 163). SFL descriptions of the genres and registers which create knowledge in school subjects have been used successfully in in-service courses for teachers working with pupils for whom the language of schooling is not their L1 (e.g. Schleppegrell, Achugar, & Otéiza, 2004). This approach to writing is completely integrated into the teaching of a topic from the curriculum (see Rose & Martin, 2012, for a detailed proposal). The process starts with scaffolded deconstruction of a model text, directing attention to its functional stages and register patterns, expanding students' resources for the types of meanings which will be required in the subsequent writing task.[5]

5.   This responds to Gentil's (2011) call for work at the level of register, since many L2 writers may learn the stages of a genre but still lack lexico-grammatical resources.

The activity of writing is also modeled, leading gradually to independent writing by students. In this process, the spoken language plays an important role, through "instructional conversation" (see Weissberg, 2006, p. 61), as 'talk around the text' uses meta-language which has been gradually brought into discussions about linguistic choices in the texts being read or written in class. Studies in different parts of the world focusing on, or including, non-L1 writers (e.g. de Oliveira & Lan, 2014; Harman, 2013; Humphrey & Macnaught, 2016; Rose, 2010; Schleppegrell et al., 2004; Whittaker & Acevedo, 2016) have shown that explicit, language-based approaches to writing in content classes can improve the quality of texts in an L2 or an FL. This is an area in which more specific and controlled research projects would be very valuable.

References

Beacco, J.-C. (2010). *Items for a description of linguistic competence in the language of schooling necessary for learning/teaching history (end of obligatory education). An approach with reference points*. Strasbourg: Council of Europe. <https://www.coe.int/en/web/platform-plurilingual-intercultural-language-education/language-s-in-other-subjects> (26 May, 2020).

Beacco, J.-C. (2017). Language in all subjects: The Council of Europe's perspective. *EuJAL 5*, 157–176. https://doi.org/10.1515/eujal-2017-0004

Byrnes, H. (2006). What kind of resource is language and why does it matter for advanced language learning. An introduction. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 1–8). London: Continuum.

Byrnes, H. (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education*, 20, 50–66. https://doi.org/10.1016/j.linged.2009.01.005

Byrnes, H. (2011). Beyond writing as language learning or content learning. Construing foreign language writing as meaning-making. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language*, (pp. 133–157). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.11byr

Christie, F. (2002). The development of abstraction in adolescence in subject English. In M. Schleppegrell & C. Colombi (Eds.), *Developing advanced literacy in first and second languages* (pp. 45–66). Mahwah, NJ: Lawrence Erlbaum Associates.

Christie, F. (2012). *Language education throughout the school years: A functional perspective*. Chichester: Wiley-Blackwell.

Christie, F., & Derewianka, B. (2008). *School discourse*. London: Continuum.

Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. London: Continuum.

Colombi, M. C. (2009). A systemic functional approach to teaching Spanish for heritage speakers in the United States. *Linguistics and Education*, 20, 39–49. https://doi.org/10.1016/j.linged.2009.01.004

Coetzee-Lachmann, D. (2007). Assessment of subject-specific task performance of bilingual geography learners: Analysing aspects of subject-specific written discourse (Unpublished doctoral dissertation). Osnabrück University.

Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms*. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.20

De Graff, R., Koopman, G. J., & Westhoff, G. (2007). Identifying effective L2 pedagogy in content and language integrated learning (CLIL). *Viewz* 16(3), 12–19.

De Oliveira, L. C. (2011). *Knowing and wrting school history: The language of students' expository writing and teacher expectations*. Charlotte, NC: Information Age Publishing.

De Oliveira, L. C., & Lan, S.-W. (2014). Writing science in an upper elementary classroom: A genre-based approach to teaching English language learners. *Journal of Second Language Writing*, 25, 23–39. https://doi.org/10.1016/j.jslw.2014.05.001

Eggins, S., Martin, J. R., & Wignell, P. (1993). The discourse of history: Distancing the recoverable past. In M. Ghadessy (Ed.), *Register analysis: Theory and practice* (pp. 75–109). London: Pinter.

European Commission/EACEA/Eurydice, (2017). *Key data on teaching languages at school in Europe: 2017 edition*. Eurydice Report. Luxembourg: Publications Office of the European Union. https://doi.org/10.2797/839825

Fontaine, L. (2017). Fluid boundaries and the categorization of nominal expressions. In S. Neumann, R. Wegener, J. Fest, P. Niemietz and N. Hützen (Eds.), *Challenging boundaries in linguistics: Systemic functional perspectives* (pp. 19–38). Frankfurt: Peter Lang.

Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015). Development of EFL writing over three years in secondary education: CLIL and non-CLIL settings. *Language Learning Journal*, 43(3), 286–303. https://doi.org/10.1080/09571736.2015.1053278

Gentil, G. (2011). A biliteracy agenda for genre research. *Journal of Second Language Writing*, 20(1), 6–23. https://doi.org/10.1016/j.jslw.2010.12.006

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education*, 5(2), 93–116. https://doi.org/10.1016/0898-5898(93)90026-7

Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's introduction to functional grammar*. London: Routledge. https://doi.org/10.4324/9780203783771

Harman, R. (2013). Literary intertextuality in genre-based pedagogies: Building lexical cohesion in fifth-grade L2 writing. *Journal of Second Language Writing*, 22, 125–140. https://doi.org/10.1016/j.jslw.2013.03.006

Heine, L. (2010). *Problem solving in a foreign language: A study in content and language integrated learning*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110224467

Hirvela, A. (2011). Writing to learn in content areas: Research insights. In R. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 37–60). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.06hir

Humphrey, S., & Macnaught, L. (2016). Functional language instruction and the writing growth of English language learners in the middle years. *TESOL Quarterly*, 50(4), 792–816. https://doi.org/10.1002/tesq.247

Järvinen, H. M. (2010). Language as a meaning making resource in learning and teaching content: Analysing historical writing in content and language integrated learning. In C. Dalton-Puffer, T. Nikula, & U. Smit, *Language use and language learning in CLIL classrooms* (pp. 145–168). Amsterdam: John Benjamins. https://doi.org/10.1075/aals.7.08jar

Jexenflicker, S., & Dalton-Puffer, C. (2010). The CLIL differential: Comparing the writing of CLIL and non-CLIL students in higher colleges of technology. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 169–89). Amsterdam: John Benjamins. https://doi.org/10.1075/aals.7.09jex

Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1, 30–41. https://doi.org/10.2174/1874913500801010030

Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching*, 5(1), 3–18. https://doi.org/10.1080/17501229.2010.519030

Liardét, C. L. (2013). An exploration of Chinese EFL learner's deployment of grammatical metaphor: Learning to make academically valued meanings. *Journal of Second Language Writing*, 22, 164–178. https://doi.org/10.1016/j.jslw.2013.03.008

Liardét, C. L. (2015). Academic literacy and grammatical metaphor: Mapping development. *TESOL International Journal*, 10(1), 29–46.

Liardét, C. L. (2016). Grammatical metaphor: Distinguishing success. *Journal of English for Academic Purposes*, 22, 109–118. https://doi.org/10.1016/j.jeap.2016.01.009

Llinares, A., & Whittaker, R. (2010). Writing and speaking in a history class: A comparative analysis of CLIL and first language contexts. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 125–144). Amsterdam: John Benjamins.

Lorenzo, F. (2016). Historical literacy in bilingual settings: Cognitive academic language in CLIL history narratives. *Linguistics and Education*, 37, 32–41. https://doi.org/10.1016/j.linged.2016.11.002

Lorenzo, F., & Dalton-Puffer, C. (2016). Historical literacy in CLIL: Telling the past in a second language. In T. Nikula, E. Dafouz, P. Moore, & U. Smit (Eds.), *Conceptualising integration in CLIL and multilingual education* (pp. 55–72). Clevedon, UK: Multilingual Matters. https://doi.org/10.21832/9781783096145-006

Lorenzo, F., & Moore, P. (2010). On the natural emergence of language structures in CLIL: Towards a theory of European educational bilingualism. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 23–38). Amsterdam: John Benjamins. https://doi.org/10.1075/aals.7.02lor

Lorenzo, F., & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System*, 47, 64–72. https://doi.org/10.1016/j.system.2014.09.016

Manchón, R. (2011). Writing to learn the language: Issues in theory and research. In R. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 61–82). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.07man

Manchón, R. M., & Roca de Larios, J. (2011). Exploring learners' perceptions of the language learning potential of L2 writing. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 181–207). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31.13man

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Martin, J. R., & White, P. (2005). *The language of evaluation: Appraisal in English*. Houndmills: Palgrave Macmillan. https://doi.org/10.1057/9780230511910

Maton, K. (2013). Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education*, 24(1), 8–22. https://doi.org/10.1016/j.linged.2012.11.005

McCabe, A., & Whittaker, R. (2017). Genre and appraisal in CLIL history texts: Developing the voice of the historian. In A. Llinares, & T. Morton (Eds.), *Applied linguistics perspectives on CLIL* (pp. 105–124). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.47.07mcc

Meyer, O., & Coyle, D. (2017). Pluriliteracies teaching for learning: Conceptualizing progression for deeper learning in literacies development. *EuJAL*, 5, 199–222.

Mohan, B., & Beckett, G. H. (2001). A functional approach to research on content-based language learning: Recasts in causal explanations. *Canadian Modern Language Review*, 58, 133–155. https://doi.org/10.3138/cmlr.58.1.133

Navés, T. (2011). How promising are the results of integrating content and langage for EFL writing and overall EFL proficiency. In Y. Ruiz de Zarobe, J. M. Sierra, & F. Gallardo del Puerto (Eds.), *Content and foreign language integrated learning: Contributions to multilingualism in European contexts* (pp. 155–186). Bern: Peter Lang.

Painter, C., Derewianka, B., & Torr, J. (2005). From microfunction to metaphor: learning language and learning through language. In R. Hasan, C. Matthiessen, & J. Webster (Eds.), *Continuing discourse on language: A functional perspective*, Vol. II (pp. 563–588). London: Equinox.

Polio, C., & Park, J. H. (2016). Language development in second language writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 287–306). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-016

Ravelli, L. (2003). Renewal of connection. Integrating theory and practice in an understanding of grammatical metaphor. In A. M. Simon-Vandenbergen, M. Taverniers, & L. Ravelli (Eds.), *Grammatical metaphor: Views from systemic functional linguistics* (pp. 37–64). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.236.04rav

Roquet, H., & Pérez-Vidal, C. (2016). Do productive skills improve in content and language integrated contexts? The case of writing. *Applied Linguistics*, 38(4), 1–24.

Rose, D. (2010). *Reading to learn: Implementation and outcomes of the professional development program*. Report for the Western NSW Region. NSW Department of Education and Training. NSW, Australia.

Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School*. London: Continuum.

Ruiz de Zarobe, Y. (2010). Written production and CLIL: An empirical study. In C. Dalton-Puffer, T. Nikula, & U. Smit, *Language use and language learning in CLIL classrooms* (pp. 191–210). Amsterdam: John Benjamins. https://doi.org/10.1075/aals.7.10rui

Ryshina-Pankova, M. (2015). A meaning-based approach to the study of complexity in L2 writing: The case of grammatical metaphor. *Journal of Second Language Writing*, 29, 51–63. https://doi.org/10.1016/j.jslw.2015.06.005

Ryshina-Pankova, M., & Byrnes, H. (2013). Writing as learning to know: Tracing knowledge construction in L2 German compositions. *Journal of Second Language Writing*, 22(2), 179–197. https://doi.org/10.1016/j.jslw.2013.03.009

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410610317

Schleppegrell, M. J. (2011). Supporting disciplinary learning through language analysis: developing historical literacy. In F. Christie & K. Maton (Eds.), *Disciplinarity: Functional linguistic and sociological perspectives* (pp. 197–216). London: Bloomsbury.

Schleppegrell, M. J., Achugar, M., & Otéiza, T. (2004). The grammar of history: enhancing content-based instruction through a functional focus on language. *TESOL Quarterly*, 38(1), 67–93. https://doi.org/10.2307/3588259

Schmid, H. J. (2000). *English abstract nouns as conceptual shells. From corpus to cognition*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110808704

Veel, R., & Coffin, C. (1996). Learning to think like an historian: the language of secondary school history. In R. Hasan & G. Williams (Eds.), *Literacy in society* (pp. 191–231). London: Longman.

Whittaker, R., & Acevedo, C. (2016). Working on literacy in CLIL/bilingual contexts: Reading to Learn and teacher development. *Estudios sobre Educacion*, 31, 37–55. https://doi.org/10.15581/004.31.37-55

Whittaker, R., & Llinares, A. (2009). CLIL in social science classrooms: Analysis of spoken and written productions. In Y. Ruiz de Zarobe & R. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 215–234). Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847691675-015

Whittaker, R., & Llinares, A. (2011). Functions of language in CLIL secondary classes: A systemic functional approach. In C. Escobar & L. Nussbaum (Eds.), *Aprendre en una altra llengua, Learning through another language, Aprender en otra lengua* (pp. 141–161). Barcelona: U.A.B.

Whittaker, R., Llinares, A., & McCabe, A. (2011). Written discourse development in CLIL at secondary school. *Language Teaching Research*, 15(3), 343–362. https://doi.org/10.1177/1362168811401154

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Weissberg, R. (2006). *Connecting speaking and writing in second language writing instruction*. Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.6702

Wolfe-Quintero, K., Inagaki, S. & Kim, H.-Y. (1998). *Second language development in writing: measures of fluency, accuracy and complexity*. Honolulu, HI: University of Hawai'i Press.

Advances in future research agendas

CHAPTER 14

An ISLA perspective on L2 learning through writing

Implications for future research agendas

Rosa M. Manchón and Ronald P. Leow

University of Murcia / Georgetown University

This chapter argues for the relevance of framing L2 writing research associated with language learning as part of a *language curriculum* within an instructed second language acquisition (ISLA) perspective. In line with current ISLA theorizing, we argue that the field should prioritize the conduct of additional studies of the processing dimension of writing and we envisage this future work as contributing to both *applied ISLA* and *ISLA applied*. We synthesize past research on writing processes and, against this background, we propose several future avenues and methodological directions for a process-oriented agenda in the domain. We conclude with a synthesis of the ideas discussed and with an assessment of what our optic offers to the collective inquiry into writing as a site for language learning that this book attempts to provide.

Introduction

The acquisition of literacy is contingent upon education. Accordingly, as argued in other contributions to the book (see especially chapters in Part 1), any writing activity should be viewed within the context in which it occurs, namely, an instructed setting. Following from here, the central assumption guiding our analysis in this chapter is that when the spotlight is directed at the manner in which writing in an additional language leads to language learning as part of a *language curriculum*, the inquiry ought to be situated within an instructed second language acquisition (ISLA) framework. In adopting this optic we are concerned with one of the dimensions of the connection between writing and language learning mentioned by Cumming (Chapter 2), namely "learning *through* writing". More precisely, Cumming suggests that "L2 learning while writing can be considered either narrowly or broadly" (p. 40). From the broad perspective, the connection can be viewed through various lenses that include "learning *through, by, for,* or

with writing" (emphasis in original). Narrowly conceived, Cumming argues, "L2 learning *through* writing appears to happen *through* cognitive problem solving and restructuring, applications and enhancements of self-regulation, and collaborations with others while writing" (p. 40). Based on these distinctions, in this chapter we adopt the narrow view as our aim is to shed light exclusively on L2 learning *through* writing as part of a language curriculum, focusing primarily on the act of writing itself (given the focus of the book) and tangentially on written corrective feedback appropriation (but see Leow, Chapter 5, this volume). We believe, nevertheless, that some of the theoretical, empirical, and methodological considerations put forward in the chapter would be equally relevant for the analysis of writing as a site for language learning in other pedagogical scenarios (including those analyzed by Byrnes in Chapter 4, this volume).

The rationale behind our focus on the processing dimension in the inquiry into "L2 learning through writing", to use Cumming's terminology, will be more evident after the elaboration of the connection between L2 writing and language learning from an ISLA perspective presented in the next section. We shall discuss how we envisage such connection and we will briefly outline the research directions and methodology implications that derive from applying this ISLA lens. In line with the aims of Part 3 in the book, we subsequently present a more detailed analysis of these fruitful directions and orientations in future ISLA-oriented writing research. We conclude the chapter with a synthesis of the ideas discussed and with an assessment of what our optic offers to the collective inquiry into writing and language learning that this book attempts to provide.

The connection between L2 writing and language learning: An ISLA perspective

ISLA and language learning through writing

Instructed second language acquisition (ISLA), although in existence for over three and a half decades, has recently witnessed a subtle shift in its definition. An early definition of ISLA as "research that concentrates on how *classroom* second language acquisition takes place" (Ellis, 1990, p. vii. Emphasis in original) attempted to correctly differentiate the context (classroom) from the more naturalistic setting subsumed under second language acquisition (SLA), while a later definition highlighted the role of *instruction as an intervention* into the L2 learning process (Ellis, 2005). A more recent definition (e.g., Loewen, 2015) has gone one step further and put greater emphasis on (a) research that underscores an important role of a better understanding of the *cognitive processes* employed by additional language (L2)

learners while interacting with L2 data, and (b) how instruction may mediate such processing. Thus, from a synthesis of previous definitions, Loewen (2015) defines ISLA as "a theoretically and empirically based field of academic inquiry that aims to understand *how* (authors' italics) the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one's own" (p. 2). As can be observed, this definition underscores (a) the instructed or classroom setting; (b) a focus on the "mechanisms of learning" (cognitive processes) employed in this instructed setting (which include "the processing and internalization of L2 input; the consolidation and storage of L2 knowledge, and the production of L2 output" [Loewen & Sato, 2017, p. 3]); and (c) the potential manipulation of these processes by instructional intervention to promote superior L2 learning. This focus on "cognitive processes" and "processing" parallels the cognitive processing that is criterial to Cumming's "learning through writing" referred to above.

In an effort to obtain insights into the role that cognitive processes play in L2 learning, several recent studies (e.g., Adrada-Rafael, 2017; Martin, Niu, & Leow, 2019; See Leow, 2019a for several other recent works) have explored the implementation of cognitive processes and their manipulation during task completion (mostly controlled problem-solving or reading tasks, e.g., Cerezo, Caras, & Leow, 2016; Leow, Donate, & Gutiérrez, 2019; Leow, Cerezo, Caras, & Cruz, 2019). Cognitive processes employed during input and intake processing found to contribute to successful L2 development include activation of appropriate prior knowledge, metacognition, hypothesis testing and rule formulation, and a high depth of processing that leads to awareness at the level of understanding (Leow, 2015). In contrast, despite the otherwise abundant research on writing processes (see synthetic review below), little is known empirically about the nature and potential language effects of these learning processes in the domain of L2 writing. Accordingly, a relevant future ISLA-oriented research avenue would be to pursue a more nuanced understanding of L2 writers' processing and processes during the composing and revision phases of their writing activity (along the lines of the study reported in Chapter 10, this volume. See also López-Serrano, Roca de Larios, & Manchón, 2019) and their connection with the learning mechanisms purported to be implicated in language learning through writing (e.g. Manchón & Williams, 2016; Williams, 2012. See also Chapter 1, this volume). In turn, the empirically-based knowledge obtained in this writing process-oriented inquiry could inform classroom interventions that promote learning through writing as part of language courses. We shall come back to these future needed explorations of writing processes and processing at different points throughout the chapter.

Leow and Cerezo (2016, see also Leow, 2019b, 2019c) propose three additional considerations that need to be included within any definition of ISLA, which are

especially relevant in relation to research initiatives aimed at understanding writing as a site for L2 learning. The first is the need to acknowledge that the context of ISLA is situated *within a language curriculum*, which has specific goals, curricular information, a syllabus, and expected learning outcomes to be successfully achieved by students. The logical implication would be for future research agendas in the domain to make room for new studies situated within the language curriculum, while forming part of the language syllabus students are following during the academic year (as recently done, for instance, in Caras, 2019; Coyle, Cánovas-Guirao, & Roca de Larios, 2018, as further elaborated in a later section). Approaching the study of writing from this ISLA lens would entail the posing of new research questions and, as a result, the adoption of diverse methodological procedures, the most relevant ones being to move from laboratory settings to the real language/writing classroom, to situate research within the syllabus, and to adopt a longitudinal perspective in recognition of the temporal nature of a language curriculum. These three dimensions will be further elaborated upon below when suggesting future methodological directions.

Leow and Cerezo (2016) also propose the need to seriously consider the type of learning that takes place in ISLA as a second key consideration associated with any definition of ISLA. Given the formality of instructed settings that logically leads to explicit processing and the statistical superiority of explicit or intentional over implicit or incidental learning (Leow, 2018), it seems logical to suggest that the investigation of explicit learning should be prioritized in this setting (see also Leow, Chapter 5, this volume), hence the above mentioned relevance of further research on the processing dimension of the act of writing itself, in diverse environments and for diverse purposes.

The third consideration in any definition of ISLA proposed by Leow and Cerezo (2016) addresses the need for pedagogical implications for the instructed L2 environment to be academically satisfactory given its curricular status. Consequently, ISLA-oriented L2 writing research that seeks to provide pedagogical implications needs to consider the usefulness of its findings in relation to curricular learning outcomes of the instructed setting (Leow, 2019c). In this respect, Leow (2019b) goes one step further in the characterization of ISLA by differentiating the type of research performed in the strand when viewed from contextual, processing, and curricular perspectives. To this end, he divides ISLA into two sub-strands: (a) *applied ISLA*, namely, studies that investigate the many variables in the instructed setting, and (b) *ISLA applied*, which refers to studies that seek to inform pedagogical practice via pedagogical intervention. The main distinction between *applied ISLA* and *ISLA applied* lies in the former investigating the instructed setting without any specific attempt to provide pedagogical ramifications (what others would call

"basic" or "controlled/laboratory type" ISLA research), while the latter is more dependent upon studies situated within a language curriculum that seek to inform pedagogical practice in an effort to promote a level of learning that is successful from a curricular perspective (or *ISLA applied* research). We envisage future research efforts into L2 writing and language learning as impacting both these two sub-strands of ISLA, as we further discuss in the rest of the chapter.

Implications for future research on L2 writing as a site for L2 learning

Following from the above considerations, adopting an ISLA perspective brings with it implications for empirical research on writing and language learning at two global levels: The kind of questions that are worth asking for future research to be theoretically and pedagogically relevant, on the one hand, and appropriate research methodology directions to be followed in order to provide methodologically principled answers to the empirical questions in the domain, on the other. Accordingly, if we accept that an ISLA-oriented approach to investigating learning through writing entails adopting the criterial mandate in ISLA of understanding "how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one's own" (Loewen, 2015, p. 2), future theoretical and applied ISLA-oriented writing research ought to pursue developments along three main empirical routes: (a) deeper exploration of the nature and effects of the *processing dimension* of the act of writing (and the processing of feedback); (b) cross-sectional and longitudinal investigation of the nature and effects of the *manipulation of writing processes* purported and/or found to be conducive to language learning; and (c) classroom-based and controlled laboratory-type studies aimed at ascertaining whether or not writing (and rewriting after processing feedback) leads to the kind of *language learning gains* ("development and acquisition of a language" in Loewen's definition of ISLA research) that are not only theoretically predicted, but also pedagogically expected and desirable in instructed settings. In line with arguments presented earlier, advancing research along these three routes will entail conducting research contributing to *applied ISLA* and *ISLA applied.*

   In the rest of the chapter we discuss this future process-oriented research agenda. To this end, we first provide a synthesis of past research on writing processes and, against this background, we elaborate more fully on future work intended to shed light on the language learning affordances of L2 writing. In doing so we will also comment on the manner in which the empirical studies reported in Part II fit in in our vision of needed advances in the field.

Past research on writing processes and its connection with learning through writing

In its most general formulation, the study of writing processes encompasses two different phenomena (Manchón, 2019). One corresponds to the steps in or the dynamism of text production, in both individual (e.g. Leijten, van Waes, Schiver, & Hayes, 2014) and collaborative writing conditions (e.g. Kessler, Bikowski, & Boggs, 2012). The other approach (and the one that concerns us here) corresponds to the set of cognitive operations underlying individual or collaborative written text production in pen and paper or digital environments. In this case, writing processes have been referred to as those "hidden sequences of events at the heart of L2 writers' text production" (Séror, 2013, p. 1) and, accordingly, research in the field has tried to understand "the workings of the human mind as it constructs a text" (Park & Kinginger, 2010, p. 31). This is the relevant line of research on writing processes for our present purposes and the one we briefly synthesize below. We shall focus only on those research strands that we consider to be more directly linked to the study of writing processes from a language learning angle.

A substantial amount of research efforts has been devoted to comparing writing processes in the multilingual user's total linguistic repertoire. For instance, Stevenson, Schoonen, and de Glopper (2006) compared revising behavior in L1 and L2 writing and did so with data obtained via think-aloud protocols and keystroke-logging techniques. Similarly, in her 2019 study, and making use of keystroke-logging techniques, Breuer compared L1 and L2 on-line writing behaviors throughout the three macro-writing processes (planning, writing, and revising). Adding to this cross-linguistic comparison of writing processes, a substantial body of research (using a variety – and triangulation – of methodological procedures that include eye tracking, keystroke logging, and retrospective questionnaires) has been devoted to the analysis of on-line writing behaviors, especially fluency and pausing (e.g. Ong, 2014; Révész, Kourtali, & Mazgutova, 2017; Révész, Michel, & Lee, 2017; Spelman Miller, Lindgren, & Sullivan, 2008; see also contributions to Révész & Michel, 2019a). Collectively, research insights obtained in these two strands provide robust empirical evidence of the more labor-intense nature of writing in an additional language (as manifested, for instance, in fluency and pausing behavior across languages) as well as the task-related (especially task complexity) and learner-related variables (primarily L2 proficiency) that mediate the attention to language concerns while writing, precisely the dimension of writing activity more directly linked to potential language learning gains.

Another area of intense scholarly work also highly relevant for our discussion corresponds to the study of the temporal distribution of writing processes. Roca et al. (2008), employing think-aloud data, and Gánem-Gutiérrez and Gillmore

(2018), employing a triangulation of data sources via screen capture techniques, eye tracking, and stimulated retrospective recalls, reported similar results: First, formulation, that is, converting ideas into language, was found to be the predominant process in writing conditions with and without access to sources. Second, the various activities participants engaged in did not stand an equal chance of being activated at any given time in the composing process. Both studies also reported that the cognitive activity and the distribution of writing processes across the entire writing process was proficiency-dependent. Specifically relevant for our current purposes is the first finding, that is, the predominant role of formulation in the time spent on composing for all proficiency groups. In fact, as noted in the Introduction to this volume, part of the impetus for the study of writing processes from the perspective of language learning has derived from the accumulated empirical evidence of the intense linguistic processing that characterizes writing (e.g., Cumming, 1989s, 1990; Manchón, Roca de Larios, & Murphy, 2009). This has resulted in the quest for answers to the questions of how and why such rich linguistic processing may be beneficial in terms of language development, a trend well represented in some of the chapters in this collection. Thus, in two recent studies by López-Serrano, Roca de Larios, and Manchón (2019, Chapter 10, this volume), the researchers have attempted to provide detailed descriptions of writing processes to probe into the purported nature and potential language learning affordances of the intense linguistic processing that characterizes writing. Adding to this, the study by Stiefenhöfer and Michel (Chapter 11, this volume) focused on the processing dimension of the interaction in collaborative writing, and Zalbidea (Chapter 9, this volume) investigated the nature of linguistic processing in writing as compared to speaking.

In addition to the ISLA-oriented studies reported above that have investigated cognitive processes in the composing stage and their relationship with learning, a number of more ISLA, classroom-based research have adressed a subsequent stage of the writing process by probing deeper into how L2 writers process WCF, and have done so via the use of written languaging, think aloud protocols, collaborative dialogues, and noticing charts. For example, a number of descriptive and, more recently, correlational studies have investigated WCF processing (in either individual or collaborative writing conditions) to, first, establish levels of processing and, second, assess the impact of higher and lower levels of processing of WCF on immediate revisions (e.g., Adrada-Rafael & Filgueras-Gómez, 2019; Caras, 2019; Cerezo, Manchón, & Nicolás-Conesa, 2019; Park & Kim, 2019; Suzuki, 2017). These descriptive studies represent real links with crucial issues of debate in SLA and ISLA research on the process of attention and, coincidentally, many of them are framed within Leow's (2015) recent model of learning processes in ISLA that underscores the role of depth of processing and potential levels of awareness at several processing stages (input, intake, knowledge) along the L2 learning process (see Leow,

Chapter 5, this volume, for his feedback processing framework based on his model). These works are also notable for the light they shed on the very processing of WCF, be it from the perspective of levels of noticing (Suzuki, 2017) or levels of depth of processing (Adrada-Rafael & Filgueras-Gómez, 2019; Caras, 2019; Cerezo et al., 2019; Park & Kim, 2019).

Specially worthy of comment is the fact that part of this research is representative of the much needed curricular orientation that we have advocated in previous sections, constituting also worthy attempts along the second research direction mentioned earlier, namely, the cross-sectional and longitudinal investigation of the nature and effects of the manipulation of processes purported to be conducive to language learning. Thus, Caras's (2019) exploratory study addressed both *how* learners process WCF during the revision stage of a composition and the effect of type of feedback (direct, indirect, metalinguistic, or control) on the participants' subsequent performance on the Spanish linguistic dichotomies *ser* versus *estar* and the preterit versus imperfect past tense aspects. Each participant wrote a composition (Draft 1) and revised it twice, first with the respective WCF (Draft 2), and then without (Draft 3). Crucially, the compositions written formed part of the regular assignments on the syllabus and were carefully designed to elicit the target linguistic items. In addition, the usual curricular unfocused WCF was provided while the researcher selected the focused target items for her study. Type of WCF in Caras's study had no differential effect on accuracy scores over time.

Coyle, Cánovas-Guirao, and Roca (2018) is a descriptive-interventionist study conducted in two intact primary school classrooms that explored the language learning potential of WCF (models) processing by children writing collaboratively over a period of 5 months, during which time one of the groups received training in the use of models. The research intended to shed light on three main concerns. First, to gain deeper knowledge into WCF processing by looking in detail (combining a complex product-process analysis) at the full trajectories the children followed from their initial, joint problem-solving activity while producing their texts, to their collaborative analysis and appropriation of the model texts provided, to their collaborative effort to revise their initial texts revisions on the basis of their processing and appropriation of the input provided in the model. Second, the researchers tried to ascertain the impact of the trajectories identified on the children's L2 development, which entailed an equally complex and novel analysis. The final aim of the study was to look into the role of instruction in the children's processing and appropriation of the WCF provided, a most pedagogically relevant concern in any ISLA-oriented inquiry. The study represents a methodological and an empirical contribution to extant research. Methodologically, it convincingly draws our attention to the underlying assumptions behind our central constructs,

especially the importance of rejecting a blind extrapolation of constructs developed for investigating writing processes by adults (often with a background in language and linguistics) to other populations. The implication is to be cognizant of the need to problematize this dimension in future studies, a key concern we shall come back to below. Empirically, the study is a worthy contribution to ISLA-oriented inquiries into the processing dimension of WCF processing due to the population studied (children), the clear picture provided on the complexities involved in WCF appropriation (which require further investigation), the light shed on the connection between WCF appropriation and language development, and the role of instruction in how children appropriate and benefit from receiving WCF in the form of models.

Despite these notable attempts, given the paucity of online studies that have gathered concurrent/on-line data on learner processes as they engage in both the composing and revising stages of the L2 writing process, the need to gather further process data during these two stages is clearly warranted. Additionally, given that research on writing processes from the perspective of their language learning affordances is still in its infancy, preliminary conclusions on the connection between the implementation of specific writing processes and language development constitute at present empirical questions in need of further exploration. In the next section, we suggest how to move forward in research agendas on writing processes from the perspective of language learning in terms of avenues of research and research methodology considerations.

Future research directions

Research avenues

We suggested in the opening section that relevant future ISLA-oriented research agendas ought to pursue a more nuanced understanding of L2 writers' processing and processes during the composing and revision phases of their writing activity. We further suggested that advancements would also derive from cross-sectional and longitudinal analyses of potential effects of the manipulation of processes purported to be conducive to language learning. We would now like to suggest that this future research should aim at being comprehensive. As noted by Cumming (2016, p. 65), "L2 writing is a complex, multifaceted, and variable phenomenon, realized in diverse ways by differing populations of learners producing differing kinds of texts in differing societal contexts and acted upon for differing purposes in particular educational, settlement, or workplace programs around the world". As a result, we recommend that worthy avenues to explore in the analysis of language learning affordances of writing *as part of a language curriculum* include the following:

Avenue 1: Expand current writing process-oriented research. Along the lines of some of the recent studies mentioned above that have shed light on writing processes and strategies (e.g., Coyle et al., 2018; López Serrano, Roca de Larios, & Manchón, 2019. See also studies by López Serrano et al., Stiefenhöfer & Michel, and Zalbidea, this volume), our position is that there is need for further descriptive, exploratory studies of writing processes, not only to provide more detailed data on processes and strategies employed during writing (at one particular point in time and across time, in diverse contexts, and by diverse populations, as noted below), but also to probe deeper into their language learning affordances. Such descriptive studies would eventually allow us to move from hypothesis generating to hypothesis testing research in this domain. Of relevance, almost three decades ago, in a much quoted pioneering and seminal paper, Cumming (1990) outlined the psycholinguistic rationale for the language learning potential of L2 writing in terms of writing processes implemented when attempting to produce demanding texts. A dimension in Cumming's thinking that has importance (and has not been particularly researched) was his claim for the relevance of conducting descriptive, exploratory studies of writing processes anticipated to "have potential for learning of the language" (p. 484) as a necessary, preliminary step preceding the setting up of more controlled research agendas leading to studies that would be supportive (or not) of theoretical predictions. As discussed in previous sections, the theoretical and empirical interest in writing as a site for language learning has certainly grown since. Similarly, solid empirical studies on writing processes have been conducted. Yet, the study of writing processes from the perspective of language learning is rather limited despite the fact that Cumming's suggestions about their relevance were put forward 30 years ago, hence our conviction that the study of writing processes from the perspective of their language learning potential merits further attention. Due to the scope of the chapter, we shall not discuss worthy future studies on the processing dimension of written feedback appropriation (but see Leow, Chapter 5 in the present volume).

In order to move forward in the study of the learning affordances of writing processes, future work needs to acknowledge that traditional modes of writing (i.e. paper-and-pen) are currently being shared with increasingly digital modalities. As noted by Hort (2017, quoting McKee & DeVoss, 2007), "Writing is becoming more and more digital and that development should therefore also shape writing research: The questions asked; the sites studied; the methodologies put to use." (p. 1). And she adds: "Digital technologies add up even more ways in which and places where people can write. As a consequence, research also has to find new ways to capture this process." (p. 1). We would subscribe to these suggestions for the study of the connection between writing processes and L2 learning, adding that, as Stiefenhöher and Michel (Chapter 11, this volume) suggest, research on

outside-school digital writing practices needs to be conducted. Also of relevance, and adding another layer of complexity to future writing process-oriented studies, an important limitation of purely cognitive models of writing (and hence of research intended to test or apply tenets in such models) is that they tend to overlook that, as a rule, writing activities are distributed over time (hours, days, weeks, and months). Such a limited perspective would not allow for a full understanding of the impact that time and space may play in writing and in the implementation of writing processes. Accordingly, future research on writing processes must include not only the exploration of both print-based pen-and-paper and screen-based writing processes, but also the time and space distributed nature of writing and writing processes (see Manchón, forthcoming).

Avenue 2: Expand research on task-related concerns. Another worthy avenue for future research lies in the elucidation of the role of task-related variables (most notably task modality and task complexity) in language learning through writing. Additionally, in line with the ISLA research mandate of understanding both L2 cognitive processes and how instruction may mediate such processing, it would be pedagogically relevant to expand current knowledge on the way in which the external manipulation of task variables may influence the implementation of writing processes potentially conducive to language learning.

Compatible with this ISLA perspective, and adding to past task-modality studies (see review in Manchón & Vasylets, 2019. See also Manchón, 2014), future developments ought to continue the inquiry into whether or not the differential time nature of speaking and writing may bring about different language learning effects. The available empirical work on *task modality* effects (see review in Manchón & Vasylets, 2019; see also Sánchez et al., Chapter 6, Zabildea, Chapter 9, and Vasylets, Gilabert, & Manchón, Chapter 8, this volume) distinctively points to a greater learning potential of writing over speaking tasks as assessed by the characteristics of the output produced (although conflicting findings exist regarding specific dimensions of performance, most notably in the area of syntax). Task-modality research has also shown that modality plays a crucial role in how L2 users experience tasks and in their motivation (Cho, 2018). Rather relevant from a language learning perspective is the reiteration in the extant research that speaking and writing may in fact contribute differentially to language learning. Thus, Zalbidea (2017, see also Chapter 9 for her contribution to this volume) categorically concluded that "speaking and writing tasks have the potential to direct learners' attention to the improvement of different dimensions of L2 output" (p. 349). Along similar lines, Vasylets, Gilabert, and Manchón (2017) interpreted their findings as "empirical evidence for the theoretical prediction that the oral and written modes offer different opportunities for language practice and development, and, consequently, mode can constitute a task design feature that may contribute to the development of distinct L2 competencies"

(p. 25). They further argued that the "oral and written modes may promote different kinds of L2 production and, by extension, may advance L2 competences in different but complementary ways" (p. 422).

Yet, task modality studies have for the most part been concerned with measuring effects on performance. We would therefore argue that future task-modality studies ought to add a processing dimension (more problematic in the oral mode, we admit) in an effort to shed new light on the relationship between the final written text and the processes involved in generating it. This would allow us to elucidate if and how the processes found to be conducive to language learning mentioned in an earlier section (such as attention, depth of processing, activation of appropriate prior knowledge, metacognition, or hypothesis testing and rule formulation) are linked to the superior performance observed in the written modality. Similarly, there is a clear need for future studies to probe deeper into the role of modality by, for example, testing and comparing empirically affordances and benefits in studies of grammatical and lexical development and type of learning (see also Schmitt, Chapter 15, and Polio, Chapter 16, present volume) via speaking and/or writing.

Equally relevant from empirical applied perspectives would be to expand research on the way in which the manipulation of *task complexity* may mediate the implementation of writing processes. In this respect, some of the process studies concerned with on-line writing behaviors mentioned above in the analysis of past writing processes research have provided empirical evidence of the effects of task-complexity variables on attentional resources and, hence, on the availability of attention to address language-related concerns while writing. For example, Révész et al. (2017) concluded that the less complex task in their study "reduced processing burden on planning processes, facilitating attention to linguistic encoding" (p. 208). Similarly, Ong (2014) found an effect for the manipulation of task conditions on metacognitive processes, as well as trade-off effects between metacognitive processes related to idea generation and information organization, on the one hand, and language-related dimensions of writing activity, on the other. Importantly, in their review of task complexity studies, Manchón and Vasylets (2019) conclude that the available empirical insights not only support the view that "the written mode can potentially channel task complexity effects in a somewhat more productive way than the oral mode" (p. 349), but also that it provides "evidence, albeit tentative, that the language learning potential of L2 writing tasks could be enhanced even more through the adjusted manipulation of cognitive task complexity, inducing deeper levels of processing and more effective linguistic behaviours in complex writing tasks" (p. 349). Testing these predictions would constitute another worthy avenue for future research on the connection between writing and language learning.

In short, along the lines of some of the studies reported in Part II, relevant items in future ISLA-oriented research agendas on writing as a site for language learning include the way in which writing processes and their manipulation are or can be mediated by (the manipulation of) task-related variables.

Research methodology considerations

An ISLA-oriented research agenda on L2 writing as a site for language learning poses diverse methodological challenges, of which we suggest the following should be prioritized in future empirical work:

1. The field would benefit from some *refinement of constructs and analytical tools*. Such refinements derive in part from the gradual development of research efforts and also from the gradual expansion of contexts and populations (further elaborated upon below), which logically leads to an adaptation of constructs to the new context or population selected for study. As an example, and marking a somewhat turning point in the otherwise abundant research on WCF processing, Coyle et al. (2018) rightly observed that the analysis of WCF appropriation needs to be more fine-tuned and go beyond simple binary distinctions. In this respect, they suggest that additional analytical categories should be established that allow for both the description of noticing of WCF and the outcomes of such noticing as a question of degree rather than as consummate categories. The same would apply, we would contend, to the analysis of the purported noticing processes (and associated learning outcomes) that are criterial to the demanding meaning-making nature of many forms of writing: The field needs to go beyond binary distinctions of noticing with/without understanding and perhaps move along the multiplicity of potential processes encapsulated in the concept of "depth of processing", as further discussed in the contribution to the book by Leow (Chapter 2) and Manchón (Chapter 17).

2. Work in the domain additionally needs to *expand methodological approaches* in order to capture writing processes in pen-and-paper and digital writing, as argued above. It is true that research on L1 and L2 writing processes has made use of a variety of methodological approaches (Manchón, 2019, forthcoming; Manchón & Roca, forthcoming), that include survey data (via questionnaires, interviews or process logs), verbally mediated data (i.e. concurrent/retrospective verbalizations via concurrent think-aloud protocols, stimulated recalls and, more recently, written languaging), direct observation of writing activity (via video recording and digital screen capture software applications), and on-line record of eyes movements, keyboard use, cursor movements, and mouse clicks and hand movements (via keystroke and handwriting logging programs, or eye tracking technology). Yet, as repeatedly emphasized in the relevant literature, all the above listed instruments and techniques have advantages and drawbacks, and both ought to be seriously considered when approaching the study of writing processes from the perspective of their language learning potential. In other words, the validity of instruments in relation to the research questions being posed must be a major consideration. For instance, although a common advantage of techniques such as video recording, digital screen capture, keystroke logging, and eye tracking is that they are all are unobtrusive and hence more

ecologically valid and less prone to reactivity than, for instance, think-aloud protocols (due to the lesser interference with the composing process. See Polio & Freedman, 2017), questions have been raised as to whether these methodological procedures provide traces of cognitive activity (what Galbraith & Baaijen, 2019 call "the problem of alignment". See also Galbraith & Vedder, 2019 and contributions to Lindgren & Sullivan, 2019 and to Révész & Michel, 2019a). Additionally, writing entails meaning making processes that go well beyond mere inscription (the main writing activity captured by keystroke logging techniques, such as Inputlog), hence the relevance of opting for alternative methodological procedures when we are concerned with non-digital forms of writing (the type of writing under study when using keystroke logging or screen capture techniques), or with writing activities beyond mere inscription, which could not possibly be left out in any serious quest for novel insights on learning through writing. In recognition of these issues, Révész and Michel (2019b), in their introduction to a collective volume on methodological advances in the study of writing processes, conclude that "a promising approach to overcoming the limitations of the various methods is to use multiple data-collection techniques and triangulate the data obtained, thereby increasing the likelihood that valid inferences are made (pp. 496–497). We would concur with them that this is the way to proceed, as evidenced in Stiefenhöfer and Michel's study reported in Chapter 11, present volume, in which they used data triangulation combining text mining tools, eye-tracking, and stimulated recall interviews (see their research methodology conclusions). Yet, although we fully acknowledge the benefits associated with data and instrument triangulation in the analysis of writing processes in general, and from the perspective of language learning in particular, it is also our view that the richness of concurrent think aloud protocols is unquestionable. In addition, although conducting protocol analysis is certainly time consuming and poses key challenges in terms of counteracting threats to validity, the access they provide to cognitive and mental activity during task completion is priceless.

In short, as observed by Révész and Michel (2019b, p. 492), "adopting new data-collection technologies and approaches to data analysis and combining these in innovative ways can generate new and more valid information about the L2 writing process and open up new avenues for research". Adding to this, we would like to suggest that advancements in the study of writing processes would benefit from insights in controlled studies in which the affordances of diverse instruments are tested. Considering that writing entails both output (text) production and input processing (in the form of either WCF or of diverse sources in reading-to-write tasks), it is of paramount importance to ascertain

which data collection instruments (or combination of them) are more likely to shed a stronger light on the specific processing phenomena we wished to investigate (see Coda chapter for a fuller elaboration. See also most relevant reflections in Galbraith & Vedder, 2019).

3. Progress in the study of writing as a site for language learning requires an *expansion of instructional contexts and populations*. An overview of ISLA-oriented writing process studies reveals that the majority of them have targeted academic contexts whose population is typically college-level and, in most cases, participants with a language/linguistics background. Given the likely differences between contexts and populations with respect to how L2 writers may learn through writing, future research needs to not only include other lower academic or non-academic contexts and L2 writers who do not possess much linguistic knowledge, but also perform contrastive analyses between contexts and populations to arrive at a better and more representative understanding of these differences. Similarly, our knowledge of writing processes is limited because of the overreliance on adult L2 users, a limitation that is especially problematic for the ISLA-oriented research we are advocating, one that ought to look into writing as a site for language learning in the entire diversity of L2 instructional settings in which children, adolescents, and adults perform writing as part of a language curriculum. We should also reiterate that the expansion of contexts and populations will logically lead to an adaptation of constructs to the new context or population selected for study, as recently done by Coyle et al. (2018), for instance, in their study of WCF appropriation by children.

4. Research needs to adopt a *long-term/longitudinal perspective.* In their review of the extant research on writing and language learning, Manchón and Vasylets (2019) categorically concluded that "any advancement in future research agendas is crucially dependent on whether or not a longer-term acquisitional perspective is adopted." (p. 356). In line with this claim, we argued earlier in the chapter that approaching the study of writing from an ISLA lens would entail moving from laboratory settings to the real language/writing classroom, to situate research within the syllabus, and to adopt a longitudinal perspective in recognition of the temporal nature of a language curriculum. We also advanced the ISLA relevance of conducting cross-sectional and longitudinal investigation of the nature and effects of the manipulation of processes purported to be conducive to language learning, which would entail designs intended to capture repeated opportunities for learning and development through writing (see Norris & Manchón, 2012). Again, these descriptive longitudinal studies of curriculum-long learning affordances of writing would eventually lead to hypothesis-testing interventionist studies framed within a curricular time span duration.

The much needed longitudinal dimension of future work is also justified on account of the time-distributed nature of writing: As mentioned in an earlier section, we need methodological procedures that allow us to inspect writing processes while performing writing activities that are distributed over different time spans. Seror (2013) would be an exemplary study of the kind of research we are advocating here (see also Manchón, forthcoming).

Conclusions

We have advocated at several points throughout the chapter that the analysis of writing processes (including linguistic processing while writing and the processing of WCF) may provide one of the most intriguing agendas for future investigation on whether and, if so, how, writing in an additional language can foster robust L2 learning. Coincidentally, this focus on cognitive processes aligns well with the postulations of several SLA and ISLA theoretical underpinnings (see Leow, 2019c, Chapter 5, present volume) and the recent uptick in studies probing deeper into the role of learner cognitive activity in the L2 learning process (e.g., see Leow, 2019a for several recent studies). Therefore, studying writing processes would surely contribute to strengthening SLA/ISLA-L2 writing interfaces. Additionally, this process-oriented inquiry would fit nicely into one of the global directions for research on the connection between writing and L2 learning identified by Cumming in Chapter 2 (present volume), namely, the one focused on "attention, self-regulation, knowledge consolidation, or collaboration while composing", which at micro-level Cumming suggests involves "processing levels of attention, knowledge consolidation, and self-regulation" (p. 40).

We have proposed that studies seeking to probe into potential connections between writing and L2 learning as part of a language curriculum ultimately need to be viewed from an ISLA, rather than SLA, perspective. More specifically, we have suggested that future process-oriented ISLA research ought to (1) expand its current descriptive scope in an attempt to eventually move from hypothesis generating to hypothesis testing in investigating L2 process-oriented writing as a site for language learning; (2) expand the range of writing environments (including pen-and-paper and computer-mediated writing) and writing conditions investigated (including more traditional and emergent individual and collaborative writing in diverse environments); (3) explore task modality, especially in relation to type of learning afforded by each mode (written vs. spoken) and the effect of (the manipulation of) task complexity variables on promoting learning via writing; (4) add more diverse classroom learner populations, such as children and language

learners who are not language/linguistics students; and (5) consider the frequent digital and time-distributed nature of writing. In our view, the identification of strategies, cognitive processes, and learner- and task-related variables that are associated with robust learning during writing will eventually lead to the conducting of experimental studies that seek to manipulate these variables to address their impact on L2 learning and development. Findings that reveal robust learning could then be confidently extrapolated to the instructed setting with much confidence that L2 writers' L2 development will be positively impacted by the interventions.

In terms of how to undertake this process-oriented inquiry, we have offered several methodological directions that include potential triangulation of several data elicitation procedures or techniques, underscoring the richness of data on internal processes elicited by think aloud protocols. On the basis of studies such as Coyle et al. (2018), we have also claimed that future inquiry also needs to problematize our constructs when new populations of instructed learners are investigated, and take note of what research has already hinted or shown about the complexity that characterizes the problem-solving activity of writing (and of rewriting after receiving feedback, although feedback processing has not been our focus of concern in the chapter). This complexity must be fully considered and integrated in our analytical tools and coding schemes. Finally, we have also recommended adopting long-term/longitudinal designs that would allow researchers to capture language learning through curricular, repeated, and extended opportunities to write, and address writing processes while performing writing activities that spread over different time spans. In this sense, we have suggested that future studies aiming to provide pedagogical ramifications will need to situate their designs within the language curriculum by adhering to the syllabus of the classroom, which will logically necessitate semester-long designs. Adopting a curricular perspective entails that studies, especially when conducted with the ultimate aim of informing classroom practice, go beyond one-shot designs (however theoretically-relevant the latter may be) and be conducted over a longer period of time to simulate the natural syllabus of a language class.

In short, the timely expansion of the field has afforded a host of promising research initiatives on how writing in an additional language can be a site for language learning. It is hoped that the ideas explored and suggestions made in this chapter signal fruitful empirical avenues to be explored in the future.

References

Adrada-Rafael, S. (2017). Processing the Spanish imperfect subjunctive: Depth of processing under different instructional conditions. *Applied Psycholinguistics*, 38, 477–508. https://doi.org/10.1017/S0142716416000308

Adrada-Rafael, S., & Filgueras-Gómez, M. (2019). Reactivity, language of think-aloud protocol, and depth of processing in the processing of reformulated feedback. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 201–213). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-14

Breuer, E. O. (2019). Fluency in L1 and FL writing: An analysis of planning, essay writing and final revision. In E. Lindgren & K. Sullivan (Eds.), *Observing writing. Insights from keystroke logging and handwriting* (pp. 190–211). Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004392526_010

Caras, A. (2019). Written corrective feedback in compositions and the role of depth of processing. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 188–200). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-13

Cerezo, L., Caras, A., & Leow, R. P. (2016). Effectiveness of guided induction versus deductive instruction on the development of complex Spanish "gustar" structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, 38, 265–291. https://doi.org/10.1017/S0272263116000139

Cerezo, L., Manchón, R. M., & Nicolás-Conesa, F. (2019). What do learners notice while processing written corrective feedback? A look at depth of processing via written languaging. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 173–187). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-12

Cho, M. (2018). Task complexity and modality: Exploring learners' experience from the perspective of flow. *Modern Language Journal*, 102(1), 162–180. https://doi.org/10.1111/modl.12460

Coyle, Y., Cánovas-Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, 42, 25–43. https://doi.org/10.1016/j.jslw.2018.09.002

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, 39, 81–141. https://doi.org/10.1111/j.1467-1770.1989.tb00592.x

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511. https://doi.org/10.1177/0741088390007004003

Cumming, A. (2016). Theoretical orientations to L2 writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 65–88). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-006

Ellis, R. (1990). *Instructed second language acquisition*. Cambridge, MA: Blackwell.

Ellis, R. (2005). *Instructed second language acquisition: A literature review*. Wellington, New Zealand: Ministry of Education.

Gánem-Gutiérrez, A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning*, 68(2), 469–506. https://doi.org/10.1111/lang.12280

Galbraith, D., & Baaijen, V. M. (2019). Aligning keystroke with cognitive processes in writing. In E. Lindgren & K. P. H. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 306–325). Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004392526_015

Galbraith, D. & Vedder, I. (2019). Methodological advances in investigating L2 writing processes. Challenges and perspectives. *Studies in Second Language Acquisition*, 41(3), 633–645.

Hort, S. (2017). Exploring the use of mobile technologies and process logs in writing research. *International Journal of Qualitative Methods*, 6, 1–14.

Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. *Language Learning and Technology*, 16, 91–109.

Leijten, M., Van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research*, 5, 285–337. https://doi.org/10.17239/jowr-2014.05.03.3

Leow, R. (2015). *Explicit learning in the classroom. A student-centered perspective*. London: Routledge. https://doi.org/10.4324/9781315887074

Leow, R. P. (2018). Explicit learning and depth of processing in the instructed setting: Theory, research, and practice. *Studies in English Education*, 23, 769–801. https://doi.org/10.22275/SEE.23.4.01

Leow, R. P. (Ed.). (2019a). *The Routledge handbook of second language research in classroom learning*. New York, NY: Routledge.

Leow, R. P. (2019b). From SLA > ISLA > ILL: A curricular perspective. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 485–493). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-33

Leow, R. (2019c). ISLA: How explicit or how implicit should it be? Theoretical, empirical, and pedagogical/curricular issues. *Language Teaching Research*, 23, 476–493. https://doi.org/10.1177/1362168818776674

Leow, R. P., & Cerezo, L. (2016). Deconstructing the "I" and "SLA" in ISLA: One curricular approach. *Studies in Second Language Learning and Teaching*, 6, 43–63. https://doi.org/10.14746/ssllt.2016.6.1.3

Leow, R. P., Cerezo, L., Caras, A., & Cruz, G. (2019). CALL in ISLA: Promoting depth of processing of complex L2 Spanish "Para/Por" prepositions. In R. M. DeKeyser & G. Prieto Botana (Eds.), *SLA research with implications for the classroom. Reconciling methodological demands and pedagogical applicability* (pp. 155–178). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.52.08leo

Leow, R. P., Donate, A., & Gutierrez, H. (2019). Textual enhancement and L2 development: A depth of processing perspective. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 319–332). New York, NY: Routledge. https://doi.org/10.4324/9781315165080

Lindgren, E. & Sullivan, K. (Eds.) (2019). *Observing writing. Insights from keystroke logging and handwriting*. Leiden, The Netherlans: Brill.

Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York, NY: Routledge.

Loewen, S., & Sato, M. (Eds.) (2017). Instructed second language acquisition (ISLA): An overview. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of second language acquisition* (pp. 1–12). New York, NY: Routledge. https://doi.org/10.4324/9781315676968-1

López-Serrano, S., Roca de Larios, J., & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically-motivated and empirically-based coding system. *Studies in Second Language Acquisition*, 41, 503–527. https://doi.org/10.1017/S0272263119000275

Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning – Insights from and for L2 writing* (pp. 27–52). Amsterdam: John Benjamins.

Manchón, R. M. (2019). The invisible dimension of L2 writing: Theoretical and empirical developments in research on L2 writing processes. Paper presented at the *Ethnographies of academic writing: Research and pedagogy*, Zaragoza (Spain), May.

Manchón, R. M. (Forthcoming). The contribution of ethnographically-oriented approaches to the study of L2 writing and text production processes. In I. Guillén-Galve & A. Bocanegra-Valle (Eds.), *Ethnographies of academic writing research: Theory, methods, and interpretation*. Amsterdam: John Benjamins.

Manchón, R. M. & Roca de Larios, J. (Eds.) (Forthcoming). *Research methods in the study of writing processes*. Amsterdam: John Benjamins.

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge: CUP.
https://doi.org/10.1017/9781108333603.015

Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing processes: Implications for theory. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 102–129). Bristol, UK: Multilingual Matters.
https://doi.org/10.21832/9781847691859-008

Martin, A., Niu, M., & Leow, R. P. (2019). Processing instruction, guided induction, and L2 development. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 377–392). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-26

McKee, H., & DeVoss, D. (2007). *Digital writing research: Technologies, methodologies, and ethical issues. New dimensions in computers and composition*. New York, NY: Hampton Press.

Norris, J., & Manchón, R. M. (2012). Investigating L2 writing development from multiple perspectives: Issues in theory and research. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 221–244). Berlin: De Gruyter Mouton.
https://doi.org/10.1515/9781934078303.221

Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing*, 23, 17–30.
https://doi.org/10.1016/j.jslw.2013.10.002

Park, E. S., & Kim, O. Y. (2019). Learners' engagement with indirect written corrective feedback: Depth of processing and self-correction. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 214–228). New York, NY: Routledge.
https://doi.org/10.4324/9781315165080-15

Park, K., & Kinginger, C. (2010). Writing/thinking in real time: Digital video and corpus query analysis. *Language Learning and Technology*, 14, 31–50.

Polio, C., & Freedman, D. (2017). *Understanding, evaluating and conducting second language writing research*. New York, NY: Routledge.

Révész, A., Kourtali, N., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, 67, 208–241.
https://doi.org/10.1111/lang.12205

Révész, A., & Michel, M. (Eds.). (2019a). Methodological advances in investigating L2 writing processes. Special Issue, *Studies in Second Language Acquisition*, 41(3).

Révesz, A. & Michel, M. (2019b). Introduction. *Studies in Second Language Acquisition*, 41(3), 491–501.

Révész, A., Michel, M., & Lee, M. (2017). Investigating IELTS academic Writing Task 2: Relationships between cognitive writing processes, text quality, and working memory. *IELTS Research Reports*, 3, 5–44.

Roca de Larios, J., Manchón, R. M., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behavior in the allocation of time to writing processes. *Journal of Second Language Writing*, 17, 30–47. https://doi.org/10.1016/j.jslw.2007.08.005

Séror, J. (2013). Screen capture technology: A digital window into students' writing processes. *Canadian Journal of Learning and Technology*, 39(3), 1–16.

Spelman Miller, K., Lindgren, E. & Sullivan, K. P. H. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, 42, 433–454. https://doi.org/10.1002/j.1545-7249.2008.tb00140.x

Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: A multidimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, 15, 201–233. https://doi.org/10.1016/j.jslw.2006.06.002

Suzuki, W. (2017). The effect of quality of written languaging on second language learning. *Writing & Pedagogy*, 8(3), 461–482. https://doi.org/10.1558/wap.27291

Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394–430. https://doi.org/10.1111/lang.12228

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Zalbidea, J. (2017). 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *Modern Language Journal*, 101(2), 335–352. https://doi.org/10.1111/modl.12389

Can writing facilitate the development of a richer vocabulary?

Advancing research agendas

Diane Schmitt

Nottingham Trent University

This chapter considers the extent to which existing research on vocabulary acquisition and writing adequately addresses the vocabulary learning potential of writing with the aim of advancing a research agenda for investigating the vocabulary learning potential of writing instruction. It first reviews the challenges L2 writers face with vocabulary size, word knowledge, and lexical fluency from the perspective of vocabulary research. It then considers how vocabulary is commonly operationalised in writing studies. Next, it considers how vocabulary is treated in writing instruction. This is followed by proposals for a research agenda that aims to bring the concerns of vocabulary research and writing research into closer alignment. A stronger word level focus is advocated.

Introduction

The aim of this chapter is to advance a research agenda to better understand the potential of writing and writing instruction to support L2 vocabulary development. Evidence of the need for such an agenda can be found in the dilemma posed by the following sets of research findings. First, studies have consistently demonstrated a strong positive relationship between quality ratings in writing and vocabulary use (Bulté & Housen, 2014; Kyle & Crossley, 2016). For example, Engber (1995) found a clear positive correlation between quality ratings and texts that are lexically variable and free of errors. Yet, Leki, Cumming, and Silva's (2008, p. 171) review of research on L2 writing reports that L2 writers write fewer words, exhibit a less diverse vocabulary, and make more word choice errors than their L1 counterparts. Likewise, Hinkel (2002) and Paquot's (2010) analyses of two separate learner corpora of university-level writing found that even at upper intermediate and advanced levels of proficiency, L2 writers tend to rely on a more limited repertoire of words

and phrases than L1 students writing on similar tasks. Ferris' (2006) study of written corrective feedback on lexical errors shows that even though learners do show short-term progress when responding to feedback on drafts, lexical errors remain stubbornly present in their writing over the longer term. Ferris even goes so far as to classify lexical errors as "untreatable" because there is no clear set of rules to guide teachers and students in giving and responding to feedback on lexis.

From the above research, we can categorize the range of vocabulary challenges experienced by L2 writers into three areas: (1) insufficient vocabulary size; (2) insufficient depth of knowledge for 'known' vocabulary; and (3) lack of fluency in using known and partially known vocabulary. These are all problems of inadequate L2 vocabulary acquisition. Finding a way to overcome these challenges requires developing a better understanding not only of how L2 vocabulary develops, but also how instruction can support that development.

There is a large body of L2 writing research and instructional practice that focuses on developing L2 learners' writing skills while paying only scant attention to how those learners develop the language proficiency needed to make continual improvement in their writing. However, since the mid-1990s, the situation has begun to change. Ferris (2010) chronicles the convergence of research on written corrective feedback with that of second language acquisition, while volumes edited by Manchón (2009, 2011a) demonstrate a broadening of the notion of "writing to learn" from a focus on content knowledge development to a focus on language development. This new focus on writing to learn language brings together SLA and writing research and aims to identify the potential of writing and writing instruction to promote language learning (Williams, 2012. See Manchón, Chapter 1, this volume).

Yet there is a need to further broaden the research agenda to take fuller account of vocabulary development in writing. Second language vocabulary acquisition has received less attention from SLA research than grammar acquisition (see Ellis, 2008; Ortega, 2009). Similarly, writing has received limited attention within the field of second language vocabulary research and instruction. Second language vocabulary acquisition research has traditionally focused on the development of receptive knowledge and has looked to reading as the primary driver of vocabulary learning. There is much less research focused on productive vocabulary use and even this focuses on how vocabulary knowledge supports writing development (Nation, 2013). The question of whether or to what extent writing can promote vocabulary acquisition has received far less attention. Thus, the aim of this chapter is to broaden the research focus by setting out an agenda that brings together and builds on existing research in the areas of both writing to learn language (WLL) and second language vocabulary acquisition (SLVA).

Like other chapters in this volume (see Byrnes, Chapter 4; Leow, Chapter 5, Manchón & Leow, Chapter 14), this chapter will include a strong emphasis on instruction. Much of the research on both L2 writing development and on L2 vocabulary development has been carried out under experimental conditions. Even when data has been collected in classrooms, studies very rarely report on the wider instructional context and how it might impact on the learning being measured. Schmitt's (2008, p. 339) review of instructed vocabulary learning found that the following learning conditions facilitate vocabulary acquisition:

- – increased frequency of exposure;
- – increased attention focused on the lexical item;
- – increased noticing of the lexical item;
- – increased intention to learn the lexical item;
- – a requirement to learn the lexical item (by teacher, test, syllabus);
- – a need to learn/use the lexical item (for task or for a personal goal);
- – increased manipulation of the lexical item and its properties;
- – increased amount of time spent engaging with the lexical item;
- – amount of interaction spent on the lexical item.

Consideration of the wider instructional context is key because meeting these conditions may require looking beyond a narrow focus on what happens when L2 learners write or in writing instruction on its own. Manchón and Leow (Chapter 14) highlight the need to consider what goes on in the wider curriculum to fully appreciate the contribution of writing and writing instruction to instructed second language acquisition. Future research must consider both the extent to which and how the act of writing and writing instruction in curricular contexts set up optimal learning conditions for vocabulary acquisition.

The chapter will begin with a review of how each of the three vocabulary challenges has been conceptualized in second language vocabulary research. This will be followed by a review of how vocabulary has been approached in writing research to date, considering the extent to which such research addresses the three vocabulary challenges facing L2 writers. Next it considers the extent to which vocabulary is addressed in writing instruction. The final section will consider how insights and concerns from both research perspectives can be aligned to advance a research agenda for better understanding of the potential of writing and writing instruction to support L2 vocabulary development.

Challenge one: Vocabulary size

A key feature that differentiates vocabulary acquisition from grammar acquisition is the size of the learning task, both in terms of how many words learners need to know and how much they need to know about each individual word. The most widely cited and up to date targets are Nation's (2006) estimates that in English 8–9,000 word families are needed for general, independent reading, and 6–7,000 word families for general, independent listening. What is missing from size research are production targets.

Word lists developed for pedagogic purposes attempt to provide some guidance on vocabulary size targets. However, Durrant (2016) argues that word lists and targets derived from large general corpora for researching *receptive* academic vocabulary needs may overstate how much vocabulary learners need to *produce* effective writing. Malmstrom, Pecorari, and Shaw's (2018) analysis of a corpus of successful L2 university student writing found that L2 writers can successfully convey meaning with a smaller productive vocabulary than what is needed to successfully read in the university environment. They developed a productive academic vocabulary list (PAVL – with 474 lemmas) which serves as a complement to Gardner and Davies's (2014) Academic Vocabulary List (AVL – with 3000 lemmas), which is based on texts students might read. However, Gardner and Davies (2016) argue that even though there appears to be no point at which successful student writers are required to produce all the vocabulary they encounter in their reading, it does not mean that they may not need these words for current or future productive use. They conclude that language programmes that set a smaller productive vocabulary size target may be doing their students a disservice.

Research investigating learners' actual receptive and productive vocabulary sizes generally finds that productive knowledge of the form-meaning link lags behind receptive knowledge (Schmitt, 2014). Webb (2008) shows that the gap between receptive and productive knowledge may be relatively small for the high frequency bands, but progressively widens as learners acquire words in lower frequency bands. He reported productive to receptive ratios of 88%, 73% and 65% at the 1,000-, 2,000- and 3,000-word frequency levels respectively for his Japanese learners. When Leviztky-Aviad and Laufer (2013) combined data from a cross-sectional corpus study of Israeli secondary school (6th–12th grades) and first year university student writing with data from a discrete item vocabulary size test, they found that learners' productive vocabulary size did increase as they moved through school when measured by the discrete item test. However, this was not matched by similar development in free productive use. They found no statistically significant differences in free productive use between grade levels until students reached the first year of university. These research findings confirm those of writing researchers who report

that L2 writers exhibit a limited lexical repertoire. The question for researchers is what prevents receptive vocabulary from becoming productive, particularly to the level where it can be used in free writing. The challenge for instruction is to determine what learning conditions are necessary to promote productive use of partially known words.

Challenge 2: Word knowledge

Understanding the vocabulary learning potential of L2 writing requires an appreciation of what it means to know a word and how word knowledge develops. Nation (2013) presents what is probably the most comprehensive attempt to define the construct of what it means to know a word. He characterizes word knowledge as a set of components relating to three main categories – form, meaning and use – with nine further sub-components (see Figure 1). He also distinguishes between receptive and productive knowledge of each of the sub-components. This framework highlights the complexity involved in "knowing a word" and shows that conceptualizing word knowledge as a known vs unknown dichotomy of the form-meaning

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express this meaning?
Meaning	Form and meaning	R	What meanings does this word form signal?
		P	What word form can be used to express this meaning?
	Concepts and referents	R	What is included in this concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	Grammatical functions	R	In what patterns does this word occur?
		P	In what patterns must we use this word?
	Collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use (register, frequency…)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when and how often can we use this word?

Figure 1. Nation's (2013, p. 27) word knowledge framework

link underspecifies what is entailed in word learning. A strength of this components approach is that it facilitates definitions and investigations of individual aspects of word knowledge and introduces the notion of incremental word learning. A weakness is that the list provides no indication of whether development of these aspects occurs in a linear way or whether the components develop in concert with one another (Milton & Fitzpatrick, 2014).

Nation's framework has had a significant impact on guiding the vocabulary acquisition research agenda to date. However, because of the complexity involved in word learning, researchers have struggled to fully operationalise the concept of word knowledge and to comprehensively measure learners' knowledge of the various components either individually or as an integrated whole. Instead of tackling all of Nation's nine aspects of word knowledge, most research has tended to test or track development of just one aspect at a time, e.g. collocations, word associations, or word parts (Barfield & Gyllstad, 2009; Milton & Fitzpatrick, 2014; Sasao & Webb, 2017).

A smaller number of studies have explored whether there might be a developmental hierarchy for various word knowledge components. Laufer and Goldstein (2004) investigated the form-meaning link and identified a three-level hierarchy of difficulty in terms of receptive and productive knowledge. This hierarchy implies that learners will be able to (1) *recognise* the meaning or form of a word, before they will be able to (2) *retrieve* its meaning, or (3) *retrieve* its form. Productive retrieval of word form, the ability that maps most closely to the cognitive process of accessing the lexicon when writing, appears to be the last to develop.

González-Fernández (2018) conducted a study of four components from Nation's (2013) framework with over 200 learners across a range of proficiency levels, learning environments, and two language groups – L1 Spanish and L1 Chinese. She found that the acquisition rate across the four components followed a statistically reliable implicational scale:

> Form-Meaning link meaning recognition > Collocate form recognition > Multiple-Meanings meaning recognition > Derivative form recognition > Collocate form recall > Form-Meaning link form recall > Derivative form recall > Multiple-Meanings recall   (González-Fernández & Schmitt, 2019, p. 13)

This scale indicates that recognition knowledge for all four components was mastered before any recall knowledge for those same components. The researchers conclude that the distinction between recognition and recall may be more important than any distinctions between the individual aspects of word knowledge. This accords with the findings on receptive and productive vocabulary size cited above. A key question is whether these developmental hierarchies hold when writing instruction focuses on developing explicit components of word knowledge for free productive use.

When considering the above studies, a major limitation that must be noted is that productive vocabulary use is normally operationalised at the word and sentence level. Few studies have looked at the development of word knowledge in free productive use. Laufer and Waldman (2011) and Levizky-Aviad and Laufer's (2013) cross-sectional studies of L2 writers' collocation use are an exception.

Challenge 3: Lexical fluency

Daller, Milton, and Treffers-Daller's (2007, p. 6) "lexical space" metaphor conceptualises vocabulary knowledge as three-dimensional by adding lexical fluency to the already familiar dimensions of vocabulary size and depth. Lexical fluency has been sub-divided into lexical access (or word recognition) – which is needed for receptive use – and lexical retrieval – which is needed for productive use (Snellings, van Gelderen, & de Glopper, 2004). Lexical retrieval entails encoding concepts into linguistic forms and is closely entwined with both size and word knowledge, because L2 writers can only retrieve what is available in their mental lexicon. This is true both in terms of how many lexical entries are stored and how many aspects of word knowledge are present in each entry.

Given that lexical retrieval is an essential subprocess in language production, Snellings, van Gelderen, and de Glopper (2004) propose that the speed at which writers retrieve L2 words from their lexicon is likely to impact on the availability of working memory for other writing processes. If this is the case, there is a clear benefit for vocabulary development through writing because L2 writers who can fluently retrieve most of the words in their texts would have more working memory available for the processing of new or partially known vocabulary. It seems important, then, to develop more sophisticated ways of researching fluency than simply counting the number of words produced by learners as is common in writing research (Hartshorn & Evans, 2015; Nitta & Baba, 2014).

Interim summary

Investigations of how vocabulary use in free productive writing can support vocabulary acquisition are limited. Therefore, questions about the role that writing and writing instruction might play in developing vocabulary knowledge remain unanswered. For example, is the receptive/productive gap a natural feature of development or is it due to learners having insufficient opportunities to use a wider range of the vocabulary they know in writing? Some aspects of word knowledge may only become salient for learners when they are required to write. How might

a more explicit focus on word knowledge in writing instruction impact the developmental trajectory of word knowledge? Could it disrupt the developmental scale that has been described?

Operationalising vocabulary in writing research

Much WLL research has been conducted within the purview of two existing research areas – Task-based Language Teaching (TBLT) and written corrective feedback (WCF) (see Chapter 1, this volume). Therefore, it is useful to consider how vocabulary is operationalised in each research tradition. In TBLT, the focus has been on effects of task modality, task repetition, and task complexity on the development of complexity, accuracy, and fluency (Chapters 6 to 9 in the current volume exemplify this trend). Here, vocabulary is operationalised as lexical complexity, as illustrated in Chapters 6, 7, and 8 in the current volume. Development of lexical complexity has commonly been investigated by comparing the vocabulary used in L2 learners' texts written at different points in time using a range of lexical complexity measures, with lexical diversity and lexical sophistication being the most common (Kyle, 2019). Measures of lexical complexity aim to extrapolate information about an individual's overall lexical proficiency through statistical analysis of patterns of vocabulary use in specific texts (Jarvis & Daller, 2013).

In terms of shedding light on how writing might overcome the three challenges faced by L2 writers, these measures have little to offer. Lexical complexity measures do, to a limited extent, shed light on writers' L2 vocabulary size, because many of them serve as a proxy for a more direct size measure. For example, using a varied range of vocabulary in a text requires knowing a wide range of words, and using sophisticated words requires using lower frequency vocabulary, which is normally acquired after higher frequency vocabulary (Schmitt, Schmitt, & Clapham, 2001). However, because these measures consider all the vocabulary used in a piece of writing, they tell us little or nothing about how L2 writers acquire individual words and phrases, which is key to understanding how size develops. While comparisons of the total word count of L2 writers' texts do provide some insight on the development of fluency, WLL studies investigating lexical complexity tell us nothing about the development of word knowledge.

A major concern with the use of these lexical complexity measures is construct validity. Jarvis (2013a, p. 17) states that "for most existing measures of lexical diversity, the underlying construct is essentially just the equation that is used to calculate the index." Unfortunately, this is also true for other measures of lexical complexity. Jarvis (2013b) relates the history of research on measures of lexical diversity and

lexical richness and identifies problems with both competing definitions and "terminological drift" resulting in overlapping definitions. Current research on the development of better tools for measuring lexical complexity still tends to focus on the statistical properties of the tools rather than the constructs that underpin the measures (Kyle, 2019).

A further concern with TBLT studies (including the ones in the current volume) relates to the limited number of genres included in research designs and the assumption that more lexical complexity (as currently measured) is an indicator of increased writing or lexical proficiency. Studies with L1 and L2 writers at different levels of education indicate that the relationship between lexical complexity and writing quality varies by text type and genre (e.g. Elgort, 2017; Kyle & Crossley, 2015; Olinghouse & Wilson, 2013). Therefore, increases in lexical diversity or lexical sophistication may not be consistent indicators of more proficient vocabulary use across all text types and genres. Thus, understanding how writing might support vocabulary acquisition requires research which explores how vocabulary use contributes to writing quality across a wider range of genres than have been used in existing WLL research.

In WCF studies, the focus is on whether written corrective feedback can improve L2 writers' linguistic accuracy. Vocabulary development in WCF studies tends to be operationalised as a reduction in lexical errors. WCF focuses on identifying different types of linguistic errors and uses a range of direct and indirect methods to indicate these errors to writers. Error correction codes commonly used in teaching (Ferris, 2011) and in research (Nicolás-Conesa, Manchón, & Cerezo, 2019; Polio & Shea, 2014) include spelling, wrong word form (for derivational morphology), wrong word (incorrect meaning sense), word choice (a better alternative is available for communicating intended meaning). Because error codes represent a sample of various aspects of word knowledge, WCF studies clearly address the challenge of word knowledge. However, it must be said that these codes present a rather impoverished view of word knowledge when compared to the range of components presented in Nation's (2013) framework.

Furthermore, even though many WCF researchers precisely code lexical errors and could provide detailed counts of the different types of lexical errors that occur in L2 writers' texts, few studies report on error types in any detail. More commonly, vocabulary errors are grouped into broad categories of general linguistic errors (Chandler, 2005), lexical errors (Sachs & Polio, 2007), word choice errors (Nicolás-Conesa et al., 2019), or error-free words, t-units, or clauses (Hartshorn & Evans, 2015). Thus, although WCF studies could provide finer-grained information about the trajectory of word knowledge development as L2 writers' use of words becomes more accurate, they rarely do. Even though lexical WCF is given at the

word and phrase level, development is tracked at the text level. However, to truly understand how word knowledge develops, accuracy needs to be tracked at the word and phrase level. Thus, these study designs provide no way of testing at the word level Ferris' (2006) claim that lexical errors are untreatable.

Interim summary

The focus on error and complexity at the text level rather than the acquisition of individual words and phrases is probably the biggest difference between WLL and SLVA research. It is important that both TBLT and WCF researchers are mindful of Gardner's (2013, p. 3) observation that "vocabulary is acquired by each learner on a word-by-word basis, not as whole frequency levels, bands or tiers" as without a word level focus, studies provide limited insight into the potential of writing to promote the development of vocabulary size or word knowledge.

Writing instruction

Hinkel (2002), Paquot (2010), and Leviztky-Aviad and Laufer (2013) all suggest that inadequate writing instruction and practice may be a possible cause of L2 writers' limited use of productive vocabulary. This concern is corroborated by Folse's (2010) case study investigation of "the extent to which vocabulary is encountered or rehearsed during a typical day in an [intensive English programme]" (p. 143) at a US university. In the composition course, he found only 9 instances of an "explicit vocabulary focus" over the course of a week and no requirement by the teacher for students to pay attention to vocabulary while writing or editing their own and peers' papers. In studies where data about vocabulary development through writing has been collected in writing classrooms, researchers commonly report that no explicit teaching of vocabulary occurred (Csomay & Prades, 2018; Mazgutova & Kormos, 2015). Investigating how vocabulary is treated in writing textbooks may explain why vocabulary gets so little attention in writing classrooms. An informal review of writing textbooks indicates that at beginning levels writing textbooks place a strong emphasis on the productive use of new words with the support of guided exercises and models, but as textbooks become more advanced the explicit focus on vocabulary development gets crowded out by a focus on other writing skills. When textbook activities do focus on vocabulary, there is little evidence to suggest that this focus follows through to the actual writing task (Schmitt, 2019). This lack of attention to vocabulary in writing books is mirrored in general EFL textbooks. Brown (2011) shows that beginner to intermediate textbooks do not teach much more than

form and meaning. He found that while grammatical functions and spoken form receive varied amounts of attention depending on the textbook, the remaining six aspects in Nation's (2013) word knowledge framework, those arguably most important for appropriate and error-free word use in free productive writing, receive only minimal attention. Textbooks like Frodesen and Wald (2016) which focus explicitly on vocabulary and grammar use in writing are less common. However, research is required to determine the extent to which an explicit approach transfers to free productive use. Jones and Haywood (2004) found that after 10-weeks of explicit training in the use of formulaic sequences, learners did become more aware of them but did not learn them very well or use them in free productive use.

The apparent lack of a principled instructional approach to vocabulary development in writing instruction and teaching materials more generally means that it is unlikely that the vocabulary learning potential of writing is being fulfilled. More classroom research is needed that investigates the effect on word learning of different approaches to vocabulary instruction both in writing alone and when writing instruction is integrated into the wider curriculum.

A future research agenda

This section aims to bring the WLL and SLVA research agendas into closer alignment. Williams' (2012) and Manchón and Vasylets' (2019) overviews of WLL research identify three unique features of writing that distinguish it from oral production and offer the potential for writing to promote SLA – the availability of time, the permanence of writing, and the inherent problem-solving nature of writing and resultant depth of processing writers engage in. From an SLVA perspective, it appears to be more relevant to identify the features of writing that offer the potential to promote vocabulary development in ways that are not achieved through reading alone. This distinction ensures that consideration is given to attested differences in receptive and productive knowledge (González-Fernández & Schmitt, 2019). These features include the need to retrieve rather than simply recognise vocabulary, the greater amount of word knowledge needed to accurately and appropriately communicate rather than comprehend meaning, and the opportunity to receive explicit feedback on vocabulary use. Although there are a number of areas of overlap between the two research traditions, this section will focus on three areas of SLA – the role of input, noticing/attention, and practice – as research in these areas offers significant potential to inform instruction aimed at overcoming the challenges of vocabulary size, word knowledge, and fluency. Writing itself consists of a number of cognitive processes – "planning, composing, reflecting, monitoring and retrieving knowledge" (Manchón & Williams, 2016, p. 569) – each of which may differ in their

potential for promoting language development, hence the relevance of future work on writing processes outlined in chapters 1, 14, and 17 (this volume). A further way of framing the research agenda will be to consider where and how each of the three vocabulary challenges faced by L2 writers impacts on the writing process, another dimension of a future process-oriented research agenda worth pursuing.

Input

In WLL research, input has been characterised as the use of external resources to "fill a hole" identified in either the planning or composing stages of writing, but it is more commonly associated with feedback received on a written product and is thus related to reflection and monitoring (Williams, 2012). From an SLVA perspective, there are several limitations with this perspective of input in terms of addressing the vocabulary size challenge. First, external input, if available at all, is often limited to access to a dictionary, a peer, a teacher or a researcher. This is manifest in the task cycle below, which is commonly found in instruction and research. Here writing tasks and instruction are carried out without reference to other areas of the curriculum which may provide valuable sources of lexical input:

1. Write to a prompt
2. Receive WCF
3. Redraft
4. Write to a new prompt
5. Receive WCF
6. Redraft
7. Write to a new prompt
8. Receive WCF
9. Redraft

In this model, input is only accessed at the planning and/or formulation stages of writing if learners notice 'holes' in their lexical repertoires. Given the size of the vocabulary learning task, such an unsystematic approach to input limits the likelihood of learners making measurable gains in vocabulary size for several reasons. First, if a task has been set at the appropriate proficiency level for learners, the number of holes is likely to be small. Second, dictionary look ups have been found to disrupt the writing process (Wolfersberger, 2003) and poor dictionary skills mean look ups are often unsuccessful (Chen, 2016). Wolfersberger (2003) found that when L2 writers are faced with too many lexical holes, the composing process may be at risk of partial or even complete breakdown. Existing research does not tell us whether

there is an upper limit on the number of lexical holes an L2 writer can contend with and still achieve their meaning making goals. Individual motivation and task and time constraints may also influence whether L2 writers choose to seek outside input to fill holes or compensate by using known words that only approximate their intended meaning or rephrasing to avoid needing to engage in an external search for a new word (Albrechtsen, 2008; Swain & Lapkin, 1995). Even when writers intentionally seek out new L2 forms to convey their intended meaning, existing research designs tend not to track new vocabulary identified at the formulation stage through to the feedback and redrafting stages, so it is unclear whether the number and quality of encounters with any new vocabulary are sufficient to leave a strong enough memory trace for long-term retention.

Task 1

Research is needed to determine whether dictionary, peer, teacher, or researcher input at the formulation or feedback stages results in durable learning of new vocabulary. Coyle, Cánovas-Guirao, and Roca de Larios' (2018) study, which identifies learning trajectories across multiple stages of the writing process, offers a model for tracking vocabulary development.

A more systematic approach to input is possible if the writing cycle is expanded to include vocabulary input from a wider range of sources. However, even then, it is not enough to simply expose learners to new vocabulary for them to use it in their writing (Lee, 2003). Lee's study follows the path of specific target words from input into multiple cycles of learner output. She found limited use of target vocabulary from reading input when there was no explicit requirement to use it. However, she convincingly shows that combining reading input with an explicit focus on target vocabulary and strong encouragement to use it can push a substantial number of both partially known and new words to productive use in a single instructional cycle. This study demonstrates the value of planned input to achieving measurable growth in vocabulary size and parallels vocabulary research for reading which consistently shows that instruction that includes an explicit focus on target words is faster and more durable than incidental encounters with words (Schmitt, 2008).

Task 2

Additional studies which provide planned vocabulary input are needed to confirm Lee's findings. Designs should include different populations, different forms of input, and different genres.

Attention

Lee's study does not tell us about how well her L2 writers used the vocabulary in their essays. The explicit instruction did not go beyond introducing and elaborating on the form-meaning link, so in the case of the new vocabulary, the information in the learners' lexical entries was likely to contain only this limited word knowledge at best, but what about the entries for the vocabulary that was already known to the learners receptively? Successful lexical retrieval and use depends not just on a word being present in the mental lexicon, but also on how much detail about the word is in the lexical entry. Parry's (1993) case study of an L2 university student's acquisition of vocabulary from a university textbook provides insight into how reading contributes to populating a lexical entry. Analysis of inferred meanings of individual words from entries in the subject's notebook shows how understanding of word meaning was built up gradually. Parry noted though that the student failed to notice and interpret how the morphological structure of a word impacted both the word's meaning and its syntagmatic relationship with other words in a sentence, for example, how the *-ive* ending in the phrase "*disruptive* behaviour" signals that the behaviour causes disruption rather than simply describes it, and how the *-ant* in *informant* signals that "someone *gives* information" rather than "someone *collects* information" (p. 122). This understanding of morphology is key to accurate word use in writing. In fact, the lack of this aspect of word knowledge in a lexical entry is an example of what I believe Skehan (2009, p. 516) was referring to when he observed that lexical choice can "derail syntax[]…making it less complex and also less accurate".

In Parry's (1993) study, the student's primary goal was to understand the vocabulary well enough to learn anthropology content, so it is perhaps not surprising that her attention was not drawn to morphological form. In writing an anthropology essay though, the student would need to be aware of morphology, and many other aspects of word knowledge, to express her understanding of anthropology content. Think aloud studies of the composing processes of L2 writers demonstrate that a considerable amount of explicit attention is given to issues arising from lexical retrieval (Albrechtsen, 2008; Cumming, 1990; Manchón, Murphy, & Roca de Larios, 2007). Cumming (1990) observes that the need to attend to form-meaning relations while composing "may prompt learners to refine their linguistic expression" (p. 483) to ensure that it accurately represents their ideas and meets expected standards of usage and hypothesizes that this process of refinement may result in acquisition (see further elaboration of this position in Chapter 1, this volume). From a word knowledge perspective, it is important to explore how that refinement occurs.

Task 3

Research is needed to resolve what is essentially a chicken or the egg dilemma – does the process of solving lexical problems generate new knowledge about specific words which learners can then add to their lexical entries or does success with problem solving depend on the existing level of detail in writers' lexical entries? Existing research on lexical retrieval tends to describe and categorize the function of these episodes (Manchón, Murphy, & Roca de Larios, 2007. See also Chapter 10, this volume). However, to better understand the effect of these episodes on vocabulary acquisition, researchers need to make a closer inspection of both the episodes where learners weigh up available word choices and the effect of their ultimate choices on the success of the text, and use this information to determine whether development has occurred. This requires word level analysis of the types of word knowledge that were used during any specific evaluation and those that were lacking. To determine whether the process of composing can generate new word knowledge, evidence is needed that reflecting on existing knowledge of a specific word, on related words, or on more general knowledge about vocabulary yields new knowledge for a specific lexical entry.

Task 4

From an instructional perspective, it is important to investigate whether explicit instruction that calls attention to various aspects of word knowledge, and provides opportunities for L2 writers to build more elaborate lexical entries (Frodesen & Wald, 2016; Jullian, 2000) can lead to better quality evaluation of vocabulary choices during lexical retrieval episodes. Since age, L2 proficiency, and writing expertise can all impact on the quality of L2 writers' metalinguistic problem-solving behaviour (Albrechtsen, 2008), multiple studies that manipulate these factors are needed.

When studying the effects of written corrective feedback on word knowledge development, it is important to acknowledge that WCF's focus on accuracy is problematic for vocabulary acquisition because accuracy is an end stage, while word knowledge is a multi-dimensional construct that develops incrementally. The debate within WCF research over the value of different forms of feedback is a side issue for vocabulary as long as research and instruction follow the design illustrated at the start of this section. In this type of design, topic changes and the nature of word distribution in texts (see Jarvis, 2013a for an explanation of Zipf's law) mean that most vocabulary will receive only one shot at feedback, unlike grammar which enjoys the benefit of regular repeated feedback. It may not be the case that vocabulary is "untreatable" (Ferris, 2006), but rather that L2 writers do not benefit from repeated feedback on individual words.

Task 5

Manchón's (2011b) proposal to replace "feedback for accuracy" with "feedback for acquisition" is particularly relevant for vocabulary but requires rethinking the nature of the instruction/research cycle and of feedback. Van Beuningen, de Jong and Kuiken's (2012) experimental study comparing the effects of direct and indirect feedback on written work is an exception to the common research design, and with some slight revisions offers a model for providing incremental feedback on vocabulary. Their study was carried out with a mix of L1 and L2 Dutch secondary school students studying biology in Dutch. The (almost) exact repetition design was modelled on a classroom instructional cycle which included an input phase on the topic of metamorphosis, an email task that required students to use the biology content, feedback on writing for the experimental groups, and two post-test writing tasks. They found that learners who received indirect feedback with correction codes outperformed students in the control conditions and that indirect feedback on non-grammatical errors had the greatest long-term effect on student texts. The error coding system categorized word form as a grammatical error, while word choice, orthographic errors, and pragmatic appropriacy errors were grouped under the heading non-grammatical errors. A strength of using error correction codes in an (almost) exact task repetition design (see Manchón, 2014) embedded in a broader instructional cycle is that communicating content is the primary goal, and the functional requirements for vocabulary in the task can be identified. Word choice, word grammar and pragmatic appropriacy are constrained by the task, so lexical development can be tracked. The changes needed include providing feedback at all stages of the instructional cycle and reconsidering the range of codes available. The longitudinal studies by Amelohina, Nicolás-Conesa, and Manchón (Chapter 7, this volume) and Nicolás-Conesa et al. (2019) illustrate the trend in research design advocated here.

At the word level, the limited set of feedback codes commonly used in WCF reduces the vocabulary construct to essentially the form-meaning link. Use of a single code for word choice vastly underestimates the complexity of the "use" components in Nation's (2013) framework. Word choice is at least partly governed by topic, genre, grammar, collocation, association, and connotation. The single code offers no guidance to learners as to which of these components is the cause of any problem. These codes also fail to acknowledge the varied functions that vocabulary plays in creating effective texts. For example, lexis indicates writer stance and attitude, provides cohesion both through transition words and lexical cohesion, allows writers to repackage the ideas of others, signals group membership, and indicates domains of use. It is difficult to see how conventional error correction codes can convey to learners the level of information needed to improve vocabulary use in writing. Alternatives to codes include direct correction, models, and

reformulations, but when each piece of writing is on a different, unrelated topic, supplying correct or appropriate words only supports change within the current text or with high frequency vocabulary. Even if students do attend to a broad range of word knowledge aspects, the use of a static set of correction codes which labels word choices and word forms as simply right or wrong does not acknowledge incremental learning.

Task 6

The incremental nature of word learning makes it a good candidate for the application of "dynamic" feedback following a sociocultural approach (Lantolf & Poehner, 2005; Storch, 2018) where teachers modify their feedback to build on partial knowledge students have obtained from earlier uses. To improve the likelihood of the same words being used multiple times and available for dynamic feedback, more research is needed with tasks that approximate exact repetition designs (see Chapter 6, this volume). A word knowledge perspective also requires greater consistency in the classification of error types so that all errors across the range of word knowledge aspects are classified as vocabulary errors.

Practice

The final area of SLA to consider is practice. Although practice is commonly linked with developing lexical fluency, it makes an important contribution to all the vocabulary challenges learners face. Practice is normally operationalized in SLVA as repetition. In terms of size, research investigating the number of repetitions needed for words to be learned receptively has found numbers ranging from 6 to 20+ exposures (Schmitt, 2008). Because production requires attention to many more word knowledge components than reception, it is hypothesized that acquiring words through writing will require an even greater number of repetitions. Repetition is needed both to consolidate prior learning and to develop other word knowledge aspects. Factors affecting the number of repetitions needed include the spacing between encounters, the proficiency level of learners, the inherent properties of words, and the quality of the encounters (Schmitt, 2008). Properties of words that increase the learning burden include sound-spelling incongruence, word length, irregular morphology, similarity between lexical forms, part of speech, as well as semantic features such as the level of abstractness, level of specificity, register restrictions, idiomaticity, and multiple meaning senses (Laufer, 1997). Most of these relate to what is required for writing. To reap the benefits of incremental learning, repetitions need to be close enough together that partial knowledge is not lost to forgetting.

Three WLL studies appear to meet the goal of regular repetitions. Nitta and Baba (2014) compared the effects of exact TR (same task) and procedural TR (same task type) on the L2 development of beginning EFL writers at university. They collected 10-minute written texts once a week over two semesters, resulting in 28 data points for each learner. Because they aimed to support fluency development, they did not provide WCF. Their dynamic systems analysis suggested that exact TR drew students' attention to fluency, and procedural TR encouraged use of a wider range of words. Hartshorn and Evans' (2015) longitudinal study operationalised Skill Acquisition Theory (DeKeyser, 2007) through four principles – writing tasks and feedback should be meaningful, manageable, timely and constant. The instructional cycle included "linguistic accuracy instruction, writing practice, and the provision of dynamic WCF" (p. 18). Learners wrote 10-minute paragraphs on different topics four days per week over thirty weeks, received indirect feedback in the following lesson, kept a log of their errors and redrafted with feedback until they produced error-free texts. The indirect correction codes for vocabulary included spelling, word form, word choice and awkward wording. The effect of the treatment was measured with 30-minute timed pre- and post-tests. The results showed statistically significant gains in linguistic accuracy compared to a control group, but no meaningful differences in lexical development for eight of nine variables. Amelohina et al. (Chapter 7) investigated the effects of both exact and procedural repetition with four tasks over 24 weeks and found no change in the number of lexical errors across tasks, and a decrease in lexical sophistication.

A key strength of these studies is that students wrote regularly and over an extended period of time. Therefore, they received abundant writing practice. However, the rather disappointing results for vocabulary are not surprising, because task repetition does not ensure word repetition. Even exact task repetition with feedback appears to have limited effect at the word level, unless teachers or students set specific goals to develop vocabulary size, word knowledge, or fluency. Nitta and Baba's (2018) analysis of the reflections of both an engaged and a less engaged student from the same data set indicates that they did not set such specific goals. In fact, other writing goals appeared to divert attention away from a focus on words. The engaged student made a conscious decision to use simpler vocabulary rather than to take time to use a dictionary as a source of input. This shows a lack of awareness that enhancing the word knowledge of even "simpler" vocabulary may be necessary to use it accurately and appropriately.

Task 7

These studies illustrate the need for finer grained analyses than offered by commonly used error classification and lexical complexity measures to identify the impact of TR on the development of word knowledge. Similar to Amelohina et al.,

Horst and Collins (2006) found a decrease in lexical sophistication over time for their French L1 writers. However, reanalysis of their data at the word level showed that this was a sign of development rather than performance loss as it reflected increased use of register appropriate Anglo-Saxon words in place of lower frequency French/English cognates.

Further research is also needed to better understand how use of L1 in writing practice impacts vocabulary development. Manchón, Murphy, and Roca de Larios' (2007) review of lexical retrieval strategies reports that L2 writers regularly bypass use of external input or compensate for its lack by using the L1 in lexical retrieval. Cumming's (1990) study suggests that in such cases L2 writers use "standards of mother tongue knowledge as a reliable test of linguistic validity" (Cumming, 1990, p. 495). Qi (1998) contends that "language switching facilitates rather than inhibits L2 composing processes" (p. 429). While L1 use may be beneficial to the composing process, Jiang's (2000) psycholinguistic model of SLVA calls into question its value in developing word knowledge. This model proposes that a lexical entry in the L1 mental lexicon contains four types of highly integrated information – semantic, syntactic, morphological and formal (phonological and orthographic). Jiang suggests that instructed L2 lexical development consists of three stages of representation in the mental lexicon. In Stage 1, learners acquire only the formal aspects of the word and these serve as a pointer to the L1 entry in the mental lexicon. In Stage 2, learners have developed stronger associations between L2 word forms and L1 semantic and syntactic aspects. In Stage 3, learners have developed associations between L2 word forms and L2 semantic, syntactic and morphological aspects of words. In Stages 1 and 2, learners are still heavily reliant on L1 word knowledge for lexical processing. Increased exposure to a word means that at Stage 2 this process of association becomes automatic. The Stage 2 level of knowledge appears to be sufficient for most receptive use because associations between the L2 form and L1 word knowledge allow meaning to be accessed. Jiang suggests that for many learners their knowledge of most words does not progress beyond this stage, because there is little motivation for the levels of processing needed to extract L2 specific semantic, syntactic and morphological specifications for words. However, the lack of a full L2 specification in the mental lexicon will cause problems for production. Jiang contends that the quality of a learner's exposure to a word may impact on whether additional encounters with a word simply reinforce the association between the L2 form and L1 word knowledge or move learners towards fuller specification of L2 word knowledge.

Task 8

In studies where students are asked to write a lot, more careful attention needs to be given to the both the quantity and quality of the encounters that students have with individual words. Research designs are needed which investigate whether

writing tasks, instruction, and feedback are simply reinforcing L1 connections or developing the L2 lexical entries of individual words in ways that will support fluent, accurate and appropriate use in free productive writing.

Conclusion

Ultimately, developing a better understanding of the vocabulary learning potential of writing and writing instruction requires investigating vocabulary development at the word level rather than the text level. The first step for word learning through writing is that words need to be produced. Writing tasks need to motivate L2 writers to use new or partially known vocabulary; this addresses the size challenge. Some of this word use needs to be planned for and motivated by teachers, materials and task design to ensure that words receive multiple repetitions and occur in a range of contexts and across modalities. This is easier to achieve when writing tasks are linked to other areas of the curriculum. Learning goals should explicitly include developing both word knowledge and fluency. When these conditions have been put in place, it will be possible for researchers to investigate the contributions of different approaches to task design, genres, and feedback to the incremental development of individual words and to the wider mental lexicon.

References

Albrechtsen, D. (2008). Writing in two languages. In D. Albrechtsen, K. Haastrup, & B. Henriksen, *Vocabulary and writing in a first and second language* (pp. 112–159). Houndmills: Palgrave Macmillan. https://doi.org/10.1057/9780230593404_4

Barfield, A., & Gyllstad, H. (Eds.) (2009). *Researching collocations in another language: Multiple interpretations*. Houndmills: Palgrave Macmillan. https://doi.org/10.1057/9780230245327

Brown, D. (2011). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research*, 15(1), 83–97. https://doi.org/10.1177/1362168810383345

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Chandler, J. (2005). The efficacy of various types of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267–296. https://doi.org/10.1016/S1060-3743(03)00038-9

Chen, Y. (2016). Dictionary use for collocation production and retention: A CALL-based study. *International Journal of Lexicography*, 30(2), 225–251.

Coyle, Y., Cánovas Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, 42, 25–43. https://doi.org/10.1016/j.jslw.2018.09.002

Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes*, 33, 100–118. https://doi.org/10.1016/j.jeap.2018.02.003

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511. https://doi.org/10.1177/0741088390007004003

Daller, H., Milton, J., & Treffers-Daller, J. (Eds). (2007). *Modelling and assessing vocabulary knowledge*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511667268

DeKeyser, R. (2007). *Practice in a second language. Perspectives from applied linguistics and cognitive psychology*. Cambridge: CUP. https://doi.org/10.1017/CBO9780511667275

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, 43, 49–61. https://doi.org/10.1016/j.esp.2016.01.004

Elgort, I. (2017). Blog posts and traditional assignments by first-and second-language writers. *Language Learning and Technology*, 21(2), 52–72.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: OUP.

Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4, 139–155. https://doi.org/10.1016/1060-3743(95)90004-7

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge: CUP. https://doi.org/10.1017/CBO9781139524742.007

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA: Intersections and practical applications. *Studies in Second Language Acquisition*, 32(2), 181–201. https://doi.org/10.1017/S0272263109990490

Ferris, D. (2011). *Treatment of error in second language writing* (2nd ed.). Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.2173290

Folse, K. (2010). Is explicit vocabulary focus the reading teacher's job? *Reading in a Foreign Language*, 22(1), 139–160.

Frodesen, J., & Wald, M. (2016). *Exploring options in academic writing: Effective vocabulary and grammar use*. Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.1150364

Gardner, D. (2013). *Exploring vocabulary: Language in action*. New York, NY: Routledge. https://doi.org/10.4324/9780203798683

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305–327. https://doi.org/10.1093/applin/amt015

Gardner, D., & Davies, M. (2016). A response to "to what extent is the Academic Vocabulary List relevant to university student writing?". *English for Specific Purposes*, 43, 62–68. https://doi.org/10.1016/j.esp.2016.03.001

González-Fernández, B. (2018). Vocabulary knowledge components: Knowledge, acquisition and conceptualisation (Unpublished doctoral dissertation). University of Nottingham.

González-Fernández, B., & Schmitt, N. (2019). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics*. https://doi.org/10.1093/applin/amy057

Hartshorn, K., & Evans, N. (2015). The effects of dynamic written corrective feedback: A 30-week study. *Journal of response to writing*, 1, 6–34.

Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410602848

Horst, M., & Collins, L. (2006). From *faible* to strong: How does their vocabulary grow? *Canadian Modern Language Review*, 63(1), 83–106. https://doi.org/10.3138/cmlr.63.1.83

Jarvis, S. (2013a). Capturing the diversity in lexical diversity. *Language Learning*, 63(Supplement 1), 87–106. https://doi.org/10.1111/j.1467-9922.2012.00739.x

Jarvis, S. (2013b). Defining and measuring lexical diversity. In S. Jarvis & M. Daller (Eds), *Vocabulary knowledge: Human ratings and automated measures* (pp. 13–43). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.47.03ch1

Jarvis, S., & Daller, M. (Eds.). (2013). *Vocabulary knowledge: Human ratings and automated measures*. Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.47

Jiang, N. (2000). Lexical representation and development in a second language, *Applied Linguistics*, 21 (1), 47–77. https://doi.org/10.1093/applin/21.1.47

Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In N. Schmitt (Ed.), *Formulaic sequences*. (pp. 269–291). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.9.14jon

Jullian, P. (2000). Creating word meaning awareness. *ELT Journal*, 54, 37–46. https://doi.org/10.1093/elt/54.1.37

Kyle, K. (2019). Measuring lexical richness. In S. Webb (Ed.), *Routledge handbook of vocabulary studies*. (pp. 454–476). London: Routledge. https://doi.org/10.4324/9780429291586-29

Kyle, K., & Crossley, S. (2015). Automatically assessing lexical sophistication: Indices, tools, findings and application. *TESOL Quarterly*, 49, 757–786. https://doi.org/10.1002/tesq.194

Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, 34, 12–24. https://doi.org/10.1016/j.jslw.2016.10.003

Lantolf, J., & Poehner, M. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9, 233–265. https://doi.org/10.1191/1362168805lr166oa

Laufer, L. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds), *Vocabulary: Description, acquisition and pedagogy* (pp. 140–155). Cambridge: CUP.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength and computer adaptiveness. *Language Learning*, 54, 399–536. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learner English. *Language Learning*, 61(2), 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x

Lee, S. Y. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System*, 31, 537–561. https://doi.org/10.1016/j.system.2003.02.004

Leki, I., Cumming, A. & Silva, T. (2008). *A synthesis of research on second language writing in English*. New York, NY: Routledge.

Levizky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight year of study: Single words and collocations. *EUROSLA Monograph Series*, 2, 127–148.

Malmstrom, H., Pecorari, D. & Shaw. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes*, 50, 28–39. https://doi.org/10.1016/j.esp.2017.11.002

Manchón, R. M. (2009). *Writing in foreign language contexts. Learning, teaching, and research*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847691859

Manchón, R. M. (2011a). *Learning-to-write and writing-to-learn in an additional language.* Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.31

Manchón, R. M. (2011b). The language learning potential of writing in foreign language contexts: Lessons from research. In T. Cimasko & M. Reichelt (Eds.), *Foreign language writing instruction: Principles and practices* (pp. 44–64). Anderson, SC: Parlor Press.

Manchón, R. M. (2014). The distinctive nature of task repetition in writing. Implications for theory, research and pedagogy. *Estudios de Linguistica Inglesa Aplicada*, 14, 13–41.

Manchón, R., Murphy, L., & Roca de Larios, J. (2007). Lexical retrieval processes and strategies in second language writing: A synthesis of empirical research. *International Journal of English Studies*, 7(2), 149–174.

Manchón, R., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge: CUP. https://doi.org/10.1017/9781108333603.015

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for academic purposes programme. *Journal of Second Language Writing*, 29, 3–15. https://doi.org/10.1016/j.jslw.2015.06.004

Milton, J., & Fitzpatrick, T. (Eds.) (2014). *Dimensions of vocabulary knowledge*. Houndmills: Palgrave Macmillan. https://doi.org/10.1007/978-1-137-36831-7

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed). Cambridge: CUP. https://doi.org/10.1017/CBO9781139858656

Nicholas-Conesá, F., Manchón, R. M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new texts: Looking into feedback for accuracy and feedback for acquisition. *Modern Language Journal*, 103(4), 848–873. https://doi.org/10.1111/modl.12592

Nitta, R., & Baba, K. (2014). Task repetition and L2 writing development. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 107–136). Amsterdam: John Benjamins.

Nitta, R., & Baba, K. (2018). Understanding the benefits of repetition from a complex dynamic systems perspective: The case of a writing task. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 285–316). Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.11.11nit

Olinghouse, N., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, 26, 45–65. https://doi.org/10.1007/s11145-012-9392-5

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. London: Continuum.

Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109–129). Norwood, NJ: Ablex.

Polio, C., & Shea, M. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10–27. https://doi.org/10.1016/j.jslw.2014.09.003

Qi, D. (1998). An inquiry into language-switching in second language composing processes. *Canadian Modern Language Review*, 54, 413–435. https://doi.org/10.3138/cmlr.54.3.413

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29, 67–100. https://doi.org/10.1017/S0272263107070039

Sasao, Y., & Webb, S. (2017). The word parts levels test. *Language Teaching Research*, 21, 12–30. https://doi.org/10.1177/1362168815586083

Schmitt, D. (2019). Why don't writing textbooks reflect current vocabulary research? Paper presented at TESOL 2019 Atlanta, GA.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12, 329–363. https://doi.org/10.1177/1362168808089921

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64, 913–951. https://doi.org/10.1111/lang.12077

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88. https://doi.org/10.1177/026553220101800103

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency and lexis. *Applied Linguistics*, 30(4), 510–532. https://doi.org/10.1093/applin/amp047

Snellings, P., van Gelderen, A., & de Glopper, K. (2004). Validating a test of second language written lexical retrieval: A new measure of fluency in written language production, *Language Testing*, 21, 174–201. https://doi.org/10.1191/0265532204lt276oa

Storch, N. (2018). Written corrective feedback from sociocultural theoretical perspectives: A research agenda. *Language Teaching*, 51, 262–277. https://doi.org/10.1017/S0261444818000034

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371–391. https://doi.org/10.1093/applin/16.3.371

Van Beuningen, C., de Jong, N., & Kuiken, F. (2012). Evidence of the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62, 1–41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Webb, S. (2008). Receptive and productive vocabulary size. *Studies in Second Language Acquisition*, 30, 79–95. https://doi.org/10.1017/S0272263108080042

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Wolfersberger, M. (2003). L1 to L2 writing process and strategy transfer: A look at lower proficiency writers. *TESL-EJ*, 7. Retrieved from <http://www-writing.berkeley.edu/TESLEJ/ej26/a6.html>

Can writing facilitate the development of grammatical competence?

Advancing research agendas

Charlene Polio
Michigan State University

Although the act of writing should, theoretically, facilitate grammatical development for second language learners even more so than speaking, most of the empirical research only indirectly supports this view, and some research even contradicts it. In this chapter, I first discuss theoretical support for the role of writing in grammatical development and then summarize related research with six different foci. Next, I propose studies related to these areas, including some conceptual replications and extensions, that might help illuminate any facilitative effects, or lack thereof, on grammatical development through writing.

Introduction

Many researchers from Cumming (1990) to Harklau (2002) to Williams (2012) have posited that writing can facilitate second language (L2) acquisition. Despite the limitations of writing in promoting vocabulary learning, as explained by Schmitt (Chapter 15, this volume), some research suggests that writing, *compared to other tasks and activities* (e.g., oral communicative tasks, explicit grammar activities) promotes the retention of vocabulary. (e.g., Y. Kim, 2011; Laufer & Hulstijn, 2001; Zou, 2017). Little empirical research, however, directly shows that writing, compared to other tasks, better promotes grammatical development, nor is there much research on whether or not any learning transfers across modalities. At the same time, there is an abundance of empirical research (e.g., Barkaoui, 2016; Storch, 2010) that shows that second language learners do, to some extent, attend to grammar while writing. Overall, theoretically speaking, it is fairly clear that writing should facilitate the acquisition of L2 grammar, but because empirical research provides little direct evidence for this claim, some degree of extrapolation from the available research is required.

I begin this chapter by reviewing the discussions of why writing, even more so than speaking, should facilitate the acquisition of grammar. Next, I will discuss briefly six types of empirical studies that are related to this claim, including those that (1) examine how and to what extent writers focus on grammar as they write; (2) track grammatical development during an instructional period; (3) manipulate tasks to promote more complex, accurate, or fluent language; (4) intervene (through researcher manipulation) in writing instruction to assess the outcomes on written language learning; (5) compare grammar used by learners in writing versus speaking; and (6) study how language used in writing might transfer to speaking. In fact, only the last category of studies directly speaks to how writing might facilitate overall grammatical development but such studies are rare. The other categories of studies are, of course, related and still worthy of discussion because they provide further evidence of the potential effects of writing on grammatical development.

The second part of this chapter expands on Polio's (2017) research agenda for language development by specifically addressing how we can expand research in six areas to better understand the effects of writing on grammatical development. Embedded in these discussions are what counts as evidence of development and what we mean by writing. I see grammatical development as a broad term that refers to how learners' grammar changes, including how complex the grammar is, how much variety there is in their use of structure, and how accurately it is used. In addition, fluency can also be considered an indication of learning because it is a measure of how quickly learners can access their grammar (and lexis).

I consider both the development of implicit and explicit knowledge about language as relevant to the development of overall grammatical competence. Recall that some early L2 researchers, notably Krashen (1982, 1983), saw only unmonitored production as being relevant and, in fact, he claimed that monitored production, as done in writing, was of little benefit to oral production. This view was continued by Truscott (1996), who claimed that written corrective feedback would lead only to what he called *pseudolearning*, presumably explicit knowledge (in contrast to Manchón & Leow, this volume, who argue that explicit knowledge is a primary area of concern in understanding writing and language learning). In addition, whether implicit or explicit, we should consider how *grammar* is defined. Many of the studies discussed here, but not all, look at grammar in terms of traditional complexity and accuracy measures with a focus on grammatical form. Only a few studies mentioned (i.e., de Oliveira & Lan, 2014; Whittaker & McCabe, Chapter 13, this volume) consider a systemic functional approach that considers the relationship between form and meaning.

I also take a broad view of writing. Swain (1998) and others, for example, have studied how learners write during a dictogloss text, and a great number studies (e.g., Yoon & Polio, 2017) elicit data via five-paragraph essays similar to large scale

tests such as IELTS or the TOEFL. Byrnes (Chapter 4, this volume), on the other hand, sees writing as a type of "meaning making" (p. 73). This raises the question as to what extent students are invested in communicating meaning as they complete these more controlled writing tasks. For example, a writer might use the simplest constructions possible with the goal of avoiding errors even if a construction does not fully convey the intended meaning. Furthermore, the goal of a dictogloss is to recreate a passage, not to communicate. I agree with Byrnes that studying writing as meaning making is important, but in many language learning contexts (in contrast to writing courses), writing is used specifically to facilitate language learning (see Leow, Chapter 5, and Manchón & Leow, Chapter 14, this volume). Thus, writing in this chapter is considered as a modality applied to extended discourse (i.e., not word or sentence writing) and not necessarily (and hence not excluding) a real-life writing task, where the goal is to communicate meaning.

Why writing should facilitate grammatical development

Cumming (1990) was probably the earliest author to refer to the facilitative effects of writing on language learning. He wrote: "Composition writing elicits an attention to form-meaning relations that may prompt learners to refine their linguistic expression – and hence their control over their linguistic knowledge – so that it is more accurately representative of their thoughts and of standard usage" (p. 483). Soon after, Harklau (1992) reported on an ethnographic study of school-age learners and suggested that much of their learning took place through reading and writing, and likely more so than through listening and speaking. Meanwhile, other researchers conducted studies using writing tasks, including Swain (1998) and Swain and Lapkin (1995), and showed that students focused on language while composing (in the former, while completing a dictogloss). These early studies did not emphasize the possibility that the written modality was a factor in language learning nor were the studies framed as writing studies, perhaps because standard writing class composition-type tasks were not used.

Probably the most explicit discussion of why writing might promote the acquisition of grammar comes from Williams (2012). She explained that because writing is slower than speaking, leaves a permanent record, and requires greater precision than speaking (this last assumption challenged by Biber & Gray, 2010), it is more likely to facilitate second language acquisition (SLA). In terms of the process, she explained that writing helps learners internalize, modify, and consolidate form-meaning connections. For example, a writer might have recently learned the passive construction while reading about a procedure described in the passive, but can only understand it. The hope is that as the writer tries to use the construction, he or she may be able to check the accuracy of the grammar, perhaps getting feedback as well. The writer

may then go on to use more passive constructions with a greater variety of verbs in a greater variety of functional contexts. Admittedly, these steps are not necessarily discrete but they illustrate how writing might facilitate learning.

Williams (2012) explained SLA as a cognitive process but did not necessarily invoke any one theory of SLA; however, further support for a role for writing in SLA can be found from theories of SLA, models of writing, and what we know about the writing process. Polio (2012) detailed seven theories of SLA and explained what they would have to say about the role of written corrective feedback in language learning, but much of the discussion is relevant to all writing activities in general. In explaining how writing may contribute to grammar development, I focus briefly on what are arguably the most dominant theories of SLA: Sociocultural, Skill-Acquisition Theory, and Usage-based Theories (see Polio & Kessler, 2019, for a further discussion).

A great amount of L2 writing research has been conducted within a sociocultural framework. While much of it focuses on strategies and socialization as opposed to grammar learning (e.g., Park & De Costa, 2015), several researchers working within the framework have highlighted how writers focus on language while revising and composing. Sociocultural Theory sees learning as a social process; a learner's development is aided through the assistance of an expert or more-abled peer, which, in writing, can take place while writing collaboratively or responding to feedback. This theory has been important in explaining how contexts for writing can lead to language learning as students attend to grammatical form as they compose and revise, as illustrated in Brooks and Swain (2009) and Suzuki (2017). It seems that the grammatical scaffolding detailed in Brooks and Swain, for example, occurs more easily in writing than speaking where there is likely less focus on formulating correct language.

Skill Acquisition Theory, as described by DeKeyser (2015), views language learning as a general skill learned in the same way as other complex skills. He explains that through practice, we proceduralize this knowledge to the point that is becomes smoother and more rapid. He provides De Jong and Perfetti (2011) as one example of this in the oral modality. L2 speakers retold stories in decreasing timeframes and compared to a control group, the repetition group got faster, even on stories on a different topic. This is seen as evidence of learning. In writing research, Skill Acquisition Theory has been used to explain the effectiveness of a method of intensive written corrective feedback (e.g., Hartshorn, Evans, Merrill, Sudweeks, Strong-Krause, & Anderson, 2010; Kurzer, 2018) showing that learners were able to compose essays with fewer errors on a new piece of writing, suggesting that the intensive feedback improved grammatical competence.

Usage-based approaches maintain that large amounts of input, either oral or written, are needed for language learning and that learning of specific features

will be affected by the frequency of a feature (e.g., Ellis & Wulff, 2015); however, structures can be made more salient through feedback or explicit instruction. Any intervention that draws learners' attention to form should facilitate learning, and we have ample evidence, as discussed in the next section, that learners pay attention to form while writing, even without intervention. Writing can make grammatical structures (and vocabulary) more salient, particularly for bound morphemes that may be difficult to detect in oral speech.

All of these theories would support a facilitative role in grammatical development for any type of language production, in general, and perhaps writing even more so than speaking because of the attention that writers can and do pay to language. The role of attention can be understood by turning to Kellogg's (1996) model of the writing process, which includes planning, translating, programming, executing, reading, and editing. Johnson (2017) drew on this model to show where working memory and attention might place demands on writers. He stated, "It is likely that the translating process and monitoring system place even heavier demands on working memory capacity, as L2 writers must search long-term memory for syntactic and lexical forms to express a given idea." (p. 15). He was interested in the implications for explaining the results of task complexity research and the processes that might challenge a writer, but we can also consider the benefits that learners might incur as they muster their attentional resources during each step. Kellogg's steps seem to have parallel components in the pedagogical literature on writing: planning as prewriting; translating, programming, executing as writing; and reading and editing as revision, with or without feedback, as postwriting. As is discussed in the next section, there is ample evidence in the literature that learners pay attention to grammar at the various stages of writing.

Related areas of research

In addition to the theoretical support for a potential facilitative effect of writing on grammatical development, some empirical studies with various foci lend support to the connection between writing and grammatical development; however, some studies do not. This related research includes studies that (1) examine how and to what extent writers focus on grammar as they write; (2) track grammatical development during instruction; (3) manipulate tasks to promote more complex, accurate, or fluent language; (4) intervene in writing instruction to assess the outcomes on written language learning; (5) compare grammar used by learners in writing versus speaking; and (6) study how language used in writing might transfer to speaking. Because the focus of this chapter is on future research, this section is not exhaustive but rather a sampling of the relevant past research.

How writers focus on grammar

There is no shortage of studies showing that writers focus on grammar as they plan, write, and revise, with or without feedback. For example, Liao (2018) conducted a study of face-to-face and chat planning with learners of Chinese and found episodes where the participants focused on grammar, albeit, a fairly small number. Roca de Larios, Marín, and Murphy (2001) had students think aloud while composing and were able to identify episodes of grammar focus before learners had actually written down any text and then as they revised it. With regard to feedback, as we might expect because written feedback is by definition explicit, many studies provide evidence that learners focus on grammar as they process it, including Sachs and Polio (2007) and Suzuki (2017). And also as expected, language is often a focus as students revise collaboratively (e.g., Brooks & Swain, 2009; Storch, 2011). Whether or not this attention to language leads to the acquisition of grammar is still an open question, but it is clear that writing does indeed, to use Cumming's (1990) words, elicit "attention to form meaning relationships" (p. 483).

Although the methods for studying the writing process have expanded beyond think-aloud protocols to eye-tracking (e.g., Chukharev-Hudilainen, Saricaoglu, Torrance, & Feng, 2019) and keystroke logging (see review in Baaijen, Galbraith, & de Glopper, 2012), the writing tasks remain lab-based (but see Khuder & Harwood, 2015 and Yeh, 2014 for interesting exceptions), and I return to this methodological issue later in the chapter. Overall, we do know that learners focus, to some extent, on grammar as they write.

Written grammatical development during instruction

Byrnes (Chapter 4) makes the case for studies that track writers over the course of an educational curriculum, and Manchón and Leow (Chapter 14) also advocate curricular, longitudinal research approaches. Although not many studies can follow students over the course of an articulated curriculum, there are now many longitudinal studies of writing development in a wide variety of instructional contexts. Some show linguistic development in some areas, but the results are quite mixed. As an example, several studies have shown that for some student populations, accuracy does not increase, even up to 2.5 years of instruction (e.g., Godfrey, Treacy, & Tarone, 2014; Knoch, Rouhsahd, & Storch, 2014; Knoch, Roushad, Oon, & Storch, 2015; Roquet & Pérez-Vidal, 2017; Serrano, 2011; Storch, 2009; Yoon & Polio, 2017;) whereas four of these six studies showed an increase in fluency (but see Storch & Tapper, 2009, for a study with opposite results). While these results may seem disappointing, learners were indeed found to produce language more quickly in the context of writing and this may be evidence of learning.

Most longitudinal studies, including those just mentioned, do not directly link writing instruction to development, so it is not clear to what extent the act of writing or writing instruction does or does not facilitate learning. For example, Matzgutova and Kormos (2015) examined changes in academic written language in the context of an academic writing course and found changes related to features of academic writing, but they did not directly link instruction to language changes. Whittaker and McCabe (Chapter 13, this volume) tracked the development of grammatical metaphor (primarily nominalizations) in a four-year study of CLIL students writing about history. They specifically state that grammatical metaphor was not explicitly taught but that studying history allowed for the development of this feature as students created meaning about the course content. Of course, we do not know what a comparison group would do, particularly if explicit grammar instruction was included. Yasuda (2011) tried to more directly link language development to instruction, but direct connections were made only to formulaic sequences, which although related to grammatical competence, is not the same as the learning of new grammatical constructions.

In sum, some studies show that during writing instruction some linguistic features develop but few studies relate instruction to language improvement; thus it is difficult to determine how various instructional writing techniques or approaches might play a role in grammatical development.

The effect of task differences

Johnson (2017) recently reviewed studies examining the effects of various task variables on written language. Because of the large number of studies and variety of independent and dependent variables, I will not review them here, but these studies generally seek to understand how task features affect linguistic output, including grammar, as indicated by complexity, accuracy, and fluency measures. Somewhat ironically, the goal of these studies seems to be to determine how to best sequence tasks, presumably from easier to more difficult, as opposed to using more challenging task to elicit more complex language. Taking a different view, a recent study by Abrams (2019) found that learners of German used more complex language in an integrated reading and writing task, suggesting that such a task can push learners' language development. Similarly, Yoon and Polio (2017) noted that because students used more complex language in argumentative essays, they should not be limited to narratives simply because they might be easier. Collectively, these studies suggest writing activities that might promote grammatical development, and this point is expanded on in the next section.

Interventions

Regarding the effect of interventions at any point in the writing process, studies of written feedback are copious, and there have now been several comprehensive research syntheses and meta-analyses (e.g. Bitchener & Storch, 2016; Kang & Han, 2015) on the topic. Although short-term changes can be interesting, studies that show changes after longitudinal interventions (e.g., Hartshorn et al., 2010) are more compelling because the feedback effects could wash out quickly if other instructional and environmental factors are at play. Interventions beyond feedback that directly address language are less common.

One notable study is M. Ishikawa (2018), who found positive effects for written languaging as students compared their dictoglosses to the original passage. S. Kim (2015) found support for individualized online grammar instruction as an intervention, and Mozaffari (2017) showed that teacher-assigned pairs in a writing task led to more language-language related episodes and higher accuracy and fluency. Shintani, Aubrey, and Donnellan (2016), who also drew on Skill Acquisition Theory, found that pre- and post-task explicit instruction improved learners' use of one particular grammatical construction on a text reconstruction activity, compared to a control group, with the pre-task group doing better than the post-task. This well-constructed study is an excellent example of the kind of research that moves beyond corrective feedback, and I return to it in the next section.

Differences in spoken and written learner language

The fifth area of research relevant to the connection between writing and grammatical development compares language used in speaking versus writing, including both studies of grammatical development and studies of students' production at one point in time. Weissberg (2006) followed five adult Spanish speakers learning English over 16 weeks and found that most structures appeared first in writing for four of the participants, although there were large individual differences, and different grammatical structures behaved differently. For example, irregular verb forms appeared first in speaking but regular past appeared first in writing. This study is important because it shows that, for most learners, but not all, writing may be a place for them to try out new structures in a slower, more controlled manner.

Other studies of modality differences are not longitudinal. Manchón and Vasylets (2019), in their discussion of task modality differences, point out that learners generally use more complex language in writing than speaking, and Gilabert, Manchón, and Vasylets (2016) review studies that compare what students do in writing versus oral tasks and present an argument that writing might be better at

facilitating attention to form, as discussed earlier. Tavakoli (2014), however, found that increasing task complexity in an oral narrative task resulted in more grammatically complex language but not in a written version of the task, as did Zalbidea (2017). These findings suggest that processing constraints are different across the two modalities, but they do not necessarily show that writing can facilitate oral acquisition. Although not falling under the definition of writing as connected discourse, Zalbidea (Chapter 9, this volume) compares tasks with regard to what students focus on as they write or speak and found some evidence that writing allows for more noticing but, again, there is no direct evidence of acquisition.

Transfer across modalities

Finally, studies that examine transfer across modalities are rare but, arguably, they are most important if we are to promote writing as a way to facilitate overall grammatical competence. Liao (2018), mentioned earlier, examined how language used in oral activities might transfer to writing, but studies of how written activities might promote learning that transfers to oral activities are extremely limited. One example is Chau (2014), who had students plan for a narrative task with and without writing. Both planning groups exhibited positive changes in their language beyond a no-planning group, but there was no difference between planning with or without writing. I note here that while we can explain differences in oral and written language with regard to processing constraints, we do not have a theoretical basis for understanding the transfer or lack thereof across modes. The concept of transfer of learning has been invoked in L2 writing research (e.g., Ferris & Hayes, 2019; James, 2006) but more at the level of genre than grammar. At the same time, it is unreasonable to think that the underlying grammatical systems for writing and speaking are unrelated, but it is not clear how they might interact.

Research agendas and research tasks

The research foci discussed in the last section show areas related to evidence that writing can promote grammatical development. Yet research in each area needs to be developed and expanded to more strongly support for the claim that writing facilitates grammatical development. Drawing on the previous categories of research, in what follows I propose related research questions and detail specific studies including replication and extensions of studies that might move the field forward. There is some overlap among each of these areas, and some of the suggested studies touch upon more than one area of inquiry.

How do writers focus on grammar as they write in less-controlled writing tasks?

As stated earlier, many past studies that have described what writers do as they write have mostly focused on time-limited controlled writing tasks. This is understandable given that the research procedures used, including stimulated recall from screen captures, keystroke logging, and think alouds, are all more difficult to use outside of a lab setting. Yet, if we are to understand what writers do as they create meaning, as suggested by Byrnes (Chapter 4, this volume), we need to expand our repertoire of writing contexts, including the prompts or activities, to those that hold more investment for research participants (see also suggestions about tasks in Manchón & Leow, Chapter 14, this volume). To this end, I propose the following research tasks that will allow us to argue more broadly that learners pay attention to grammatical form while writing.

Task 1

There has been a surge of interest in multimodal writing (e.g., Belcher, 2017) Clearly, being able to communicate in multimodal platforms is an essential skill and incorporating the use of appropriate visuals along with written text can facilitate the communication of meaning. However, we do not know to what extent different types of multimodal writing affords a focus by on language, in general, and grammar in particular. Yi (2017) states that "multimodal literacy researchers do not intend to trivialize the development of language" (p. 90) but Manchón (2017) cautions that we do not know if multimodal tasks provide the same affordances for language learning as other types of writing. She says, "For multimodal writing to be valuable for language learning there must be a true demand for formulation, i.e., a real struggle to transform ideas into language, a process propitious to language development because of the learning mechanisms it would activate, and the corresponding possible changes in the L2 user's underlying linguistic system it might induce" (p. 94) Of course, not all multimodal writing tasks are the same, but including them in research on language (and grammatical) development might address Manchón's concern and move us beyond the timed essay.

One approach for future research would be to create controlled multimodal tasks and have learners think aloud as they compose or as they participate in a stimulated recall. This would allow us to evaluate the concern that multimodal tasks might not facilitate language learning. For example, students could be asked to prepare a narrated PowerPoint presentation within a certain time frame. This might determine whether or not the students were so focused on the format and visuals that they ignored linguistic problems. One issue that might arise how learners move between language used in written texts versus on slides or between oral and written language (see Cimasko & Shin, 2017, for a related case study).

Task 2

The second task is to use methods that can capture the longer process of real life writing tasks. As mentioned earlier, most of the writing process research uses, for logistical reasons, timed essays where students have no access to outside sources. Outside of testing contexts, when time pressure is less and when there is access to sources, writers likely draw on more explicit grammatical knowledge, editing strategies, dictionaries, grammar checks, corpora, and readings. If we are to suggest that writing allows for a focus on and reflection about form, we need to understand these real life processes (see also Manchón & Leow, Chapter 14, this volume). As an example, Khudar and Harwood (2015) conducted a small scale study of ten students writing an IELTS-type essay and found that some writing processes differed across timed versus untimed conditions. For example, surface-level revision was more common in the timed condition. Meaning-level revisions were more common in the untimed setting, but the global grammatical range and accuracy scores were not much different across the two conditions. This study is a good starting point to see how writing processes outside the lab could differ from those inside the lab. This study could be replicated with a larger sample size. In addition, some type of tailor-made delayed post-tests might show whether or not the grammatical revisions had any long-term impact on grammatical competence.

Case studies can also be conducted to understand how students focus (and learn) grammar in out-of-class assignments. Li and Schmitt's (2009) study of formulaic sequences in writing can serve as an example. Instead of monitoring students as they wrote, they conducted a case study in which they collected assignments over the course of year and conducted interviews with the student about where she might have learned the sequences. An alternative approach for studying grammatical development might be to have students fill out some type of log as they write outside of class noting new or challenging structures and explaining how they resolved problems.

How can or cannot writing instruction be linked to grammatical development?

Although the number of studies describing grammatical development, or lack thereof, among instructed learners are copious, and although some of the studies described curricula, they did not expand on what was happening in the classroom. We do not really know what instructional techniques or curricular are effective for facilitating grammatical development. Mixed methods research is an ideal way to better understand how language does or does not develop in writing classes because we can link previous quantitative measures with data from classroom observations, interviews, and artifacts.

Task 1

Yasuda (2011) is a good study to take as a starting point. Her mixed methods study was a no-control group semester-long study of a teaching innovation. A genre-based curriculum was implemented in an EFL class. Yasuda tracked textual features of the students' writing and included surveys and interviews to triangulate the quantitative data. The extent of her focus on language was on the use of formulaic sequences, but this study could be modified to include explicit instruction of complex structures that students had not mastered. Randomized control group studies can be logistically difficult, but the inclusion of rich qualitative data, including classroom observations, can help researchers triangulate data.

Task 2

Another approach would be to replicate other studies that tracked development or documented short-term development with some additions. For example, Mazugotova and Kormos (2015) examined how features of academic language changed over the course of a one-month intensive EAP writing class (15 hours a week plus opportunities for individual tutoring) looking at syntactic complexity and lexical development. They found vocabulary development in students at two levels of proficiency and syntactic development – in terms of noun-phrase complexity and the use of conditionals – only at the lower level. The authors state:

> The intensive nature of the EAP course, exposure to a variety of academic reading texts, individualized feedback on the overall quality of writing assignments, and immersion in the target language environment could all have contributed to increases in the lexical complexity of students' essays and to the changes observed in the use of syntactic features. Our results suggest that these instructional and environmental conditions can be conducive to lexical and syntactic development and should, if possible, be included in pre-sessional EAP programmes. It might also prove useful to call learners' attention explicitly to the lexical and syntactic features of academic writing, both before they embark on a writing assignment, for example, in the academic reading tasks they are set, and when they revise their essay.
>
> (p. 13)

Not only do the authors suggest a list of interventions that could be tested, but they also suggest that there was not an attempt made to understand why the changes took place. A conceptual replication could include observations, interviews, and learning logs to triangulate the quantitative data.

Can writing tasks be manipulated to promote grammatical development?

As discussed earlier, some tasks and genres result in more complex language although it is an open question whether this is due to some type of cognitive requirement or an affordance of the writing prompt (i.e., the task or genre requires more complex language to communicate). Additionally, no one has studied the long-term effects of these various tasks and genre on the development of grammatical complexity. If we are to argue writing promotes development, we need to show that a task that elicits more complex language leads to learning.

Task 1

An effective approach is to identify a task that has been shown to result in more complex language and test it as an intervention. For example, Johnson (2017) found a small-to-medium effect size for studies manipulating the here-and-now task complexity condition. T. Ishikawa (2007), for example, found that removing the narrative-picture prompt and telling students to write in the past tense resulted in more complex language. Johnson (2017) included only two studies of this task condition in his meta-analysis, but one potential study could have students write in these conditions over a series of tasks and then examine whether any complexity changes were transferred to a new writing task. Another approach would be to determine whether the use of more complex structures transfers to oral language either in the short-term or the long-term. For example, one could replicate T. Ishikawa's study but then have students perform the same task as an oral narrative immediately after writing to see if the same increase in complexity was maintained.

Other than feedback, what kinds of interventions
promote grammatical development?

Despite the large number of intervention studies examining the effect of feedback, we don't have evidence of what other types of interventions that might be less time-consuming might promote faster grammatical development. In fact, there are other types of interventions that hold promise based on previous studies.

Task 1

Shintani et al.'s study (2016) suggested that pre- and post-task metalinguistic instruction might be effective in promoting accuracy of one grammatical structure, namely, the past counterfactual conditional. This study is significant because it examines an efficient way to provide grammatical information that is less time consuming than corrective feedback, but we don't know how well this intervention

extends to other structures or if the instruction can transfer to similar oral tasks after writing. Thus, the study can be replicated with different grammatical structures and include a follow-up oral task.

Task 2

Another intervention that can be easily implemented but that has been rarely studied is the provision of model texts. In one study, Yang and Zhang (2010) had students look at model texts (texts written on the same prompt by native speakers) and compare them to their own writing. In their small-scale study, they looked at language-related-episodes and not outcomes. Their study could be extended to see if the use of model texts promote more complex or more accurate language.

Task 3

Related to this idea of model texts would be an intervention that focused on the grammatical features of authentic readings. Schleppegrell, Achugar, and Oteíza (2004) pointed out that in content-based language courses, there is often a focus on graphic organizers and vocabulary but not on the grammar of the content. They focused on the language of history textbooks used in middle and high school. Their focus was not on student learning but rather helping teachers focus on the language using systemic functional linguistics. Their study also offers a starting point for creating a discipline-specific intervention that could be conducted, namely, how explicit instruction on grammatical structures from the reading might promote grammatical competence.

Does written grammatical development precede oral grammatical development?

If we argue that writing pushes grammatical development, learners' written language should be, in most cases, more accurate and complex. Of course, every learner will differ in the relationship between oral and written language developments are; as noted earlier, Weissberg (2006) showed differential progress among learners in terms of which structures appeared first in the two modalities and heritage learners may be very advanced in speaking but have little or no writing ability whereas EFL learners who may have studied through reading and writing will be different. However, in some contexts, oral and written input may both be provided in roughly equal amounts, and students' oral and written language could be compared. If learners use new structures in writing before speaking, for example, this finding would lend support to the claim that writing may promote grammatical development more quickly than oral language use.

Task 1

While several studies track learners' development in oral or written language, few track both simultaneously. One exception that might help us think more about differential development across modalities is Bardovi-Harlig (2001). She examined the emergence of the present perfect in adult ESL learners collecting both oral and written data. Although she does compare the results from the oral and written tasks, clear conclusions are difficult to draw because the tasks were not directly comparable and because writing, in general, might elicit more the present perfect. Nevertheless, this study could be modified by using comparable tasks, such as oral and written narratives, collected over the course of a lower level language class, preferably in a foreign language context where input was likely more similar across participants. Possible questions might include: Do certain complex structures such as relative clauses emerge first in one modality? Do students make fewer or different errors in one modality?

Task 2

Torres (2018), studying Spanish heritage learners, conducted an intervention in which students complete a simple or a complex oral monologic task for which they received written recasts. Students in the simple condition, which did not involve reasoning demands, made more progress with the subjunctive than those in the complex condition, especially on the written assessments, which were similar to the treatment tasks. The results with regard to modality are hard to interpret since the task was oral but the feedback was written. Nevertheless, it is interesting that students did better on the written assessment. Intervention studies that include comparable oral and written tasks can be conducted to help us better understand where learning takes place first. For example, the Yang and Zhang (2010) studied mentioned earlier that had students read model texts could have included an oral narrative to be compared to the written assessment.

Can writing activities promote oral grammatical competence?

Of all the questions, this is arguably the most important if we are to test the facilitative effects of writing on overall language learning generally and the development of grammatical competence in particular. In order to better understand this, we can replicate some studies of written language with the addition of an oral task to see if learning transfers.

Task 1

One such study is Wang and Wang (2014), who examined how a written story continuation task might affect students' written language. They gave students stories in either English or Chinese, and the students had to continue them in English. Not surprisingly, students used vocabulary from the English version in their writing, but they also made fewer grammatical errors than students who had read the Chinese version. Although this study looked at the effects of reading, and not writing, we can consider whether or not any learning would last and whether it would carry over to an oral task, particularly after using any new language in the written task.

Task 2

Lynch (2001) proposed transcription of oral language as a technique for helping learners notice problems with their language. While transcription does not involve the new creation of connected discourse, students are working with their own attempts to create meaning in another modality, namely speaking. Students recorded a role play and then worked in pairs to transcribe and then revise a small segment. The teacher then reformulated the transcription and had the students review it. This article described what could be a way to harness the potential of writing to benefit oral language production. The study, however, stopped short of investigating any effect of the transcription or reformulation on subsequent language production, and this would be an obvious next step. A very small scale study conducted by Abadikhah and Valipour (2018) examined how transcription and scaffolded transcription (i.e., reviewing the transcript with an instructor) might be helpful in reducing errors in oral language. An experimental study on the learning outcomes of transcription would give more support to its usefulness.

Conclusion

Each of the different research areas contributes to our understanding of how writing can facilitate grammatical development, even if indirectly. To summarize, first, we know that learners pay attention to grammar but we need to expand the scope of inquiry to new tasks and contexts using new methods. Some of these methods, such as case studies, might even give us more insight into how specific structures were or were not learned. Second, the studies of longitudinal development in writing classes show limited progress with regard to linguistic outcomes and most do not describe actual instruction. Thus, we need studies that combine linguistic outcome measures with qualitative data that can better help us understand progress or lack thereof. Third, we know that some task and genres elicit more complex language, but we need to understand long-term effects that might facilitate grammatical

development. Fourth, although there are some small-scale studies of interventions other than feedback that appear promising, more studies need to be conducted and some need to be replicated or extended. Fifth, the relationship between oral and written language is not fully understood, likely because that relationship differs among learners. Nevertheless, we need more evidence that grammar develops first in writing for most learners if we are to maintain that writing facilitates development. Finally, and most importantly, we do not have clear evidence that structures used or learned through writing can transfer to oral language. Thus, we need to take studies of learning through writing and add an oral component to further investigate the transfer of learning across modalities.

References

Abadikah, S., & Valipour, M. (2018). The effect of self-transcription and expert scaffolding on the accuracy of oral production of EFL learners. *Pertanika Journal of Social Science & Humanities*, 26, 149–166.

Abrams, S. (2019). The effects of integrated writing on linguistic complexity in L2 writing and task complexity. *System*, 81, 110–121. https://doi.org/10.1016/j.system.2019.01.009

Baaijen, V. M., Galbraith, D. & de Glopper, K. (2012). Keystroke analysis: Reflections on procedures and measures. *Written Communication*, 29, 246–277.
https://doi.org/10.1177/0741088312451108

Bardovi-Harlig, K. (2001). Another piece of the puzzle: The emergence of the present perfect. *Language Learning*, 51(Sup), 215–264. https://doi.org/10.1111/j.1467-1770.2001.tb00018.x

Barkaoui, K. (2016). What and when second language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *Modern Language Journal*, 100, 320–340.
https://doi.org/10.1111/modl.12316

Belcher, D. (2017). On becoming facilitators of multimodal composing and digital design. *Journal of Second Language Writing*, 38, 80–85. https://doi.org/10.1016/j.jslw.2017.10.004

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9, 2–20.
https://doi.org/10.1016/j.jeap.2010.01.001

Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781783095056

Brooks, L., & Swain, M. (2009). Languaging in a collaborative setting. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 58–89). New York, NY: Routledge.

Chau, H. T. (2014). The effects of planning with writing on the fluency, complexity, and accuracy of L2 oral narratives (Unpublished doctoral dissertation). Michigan State University.

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H. H. (2019). Combined deployable keystroke logging and eye tracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, 41, 583–604. https://doi.org/10.1017/S027226311900007X

Cimasko, T., & Shin, D.-S. (2017). Multimodal resemiotization and authorial agency in an L2 writing classroom. *Written Communication*, 34, 387–413. https://doi.org/10.1177/0741088317727246

Cumming, A. (1990). Expertise in evaluating second language compositions. *Language Testing*, 7, 31–51. https://doi.org/10.1177/026553229000700104

De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 61, 533–568. https://doi.org/10.1111/j.1467-9922.2010.00620.x

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94–112). New York, NY: Routledge.

De Oliveira, L., & Lan, S. W. (2014). Writing science in an upper elementary classroom: A genre-based approach to teaching English language learners. *Journal of Second Language Writing*, 25, 23–39. https://doi.org/10.1016/j.jslw.2014.05.001

Ellis, N., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 75–93). New York, NY: Routledge.

Ferris, D., & Hayes, H. (2019). The five-paragraph essay does not transfer well into more complex types of writing. In N. Caplan & A. Johns (Eds.), *Changing practices for the L2 writing classroom* (pp. 116–132). Ann Arbor, MI: University of Michigan Press.

Gilabert, R., Manchón, R., & Vasylets, O. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, 117–135. https://doi.org/10.1017/S0267190515000112

Godfrey, L., Treacy, C., & Tarone, E. (2014). Change in French second language writing in study abroad and domestic contexts. *Foreign Language Annals*, 47, 48–65. https://doi.org/10.1111/flan.12072

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11, 329–350. https://doi.org/10.1016/S1060-3743(02)00091-7

Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, 44, 84–109. https://doi.org/10.5054/tq.2010.213781

Ishikawa, T. (2007). The effects of manipulating task complexity along the (±here-and-now) dimension on L2 written narrative discourse. In M. D. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136–156). Clevedon, UK: Multilingual Matters.

Ishikawa, M. (2018). Written languaging, learners' proficiency levels and L2 grammar learning. *System*, 74, 50–61. https://doi.org/10.1016/j.system.2018.02.017

James, M. (2006). Transfer of learning from a university content-based EAP course. *TESOL Quarterly*, 40, 783–806. https://doi.org/10.2307/40264308

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, 37, 13–38. https://doi.org/10.1016/j.jslw.2017.06.001

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, 99, 1–18. https://doi.org/10.1111/modl.12189

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Hillsdale, NJ: Lawrence Erlbaum Associates.

Khuder, B., & Harwood, N. (2015). L2 writing in test and non-test situations: Process and product. *Journal of Writing Research*, 6, 233–278. https://doi.org/10.17239/jowr-2015.06.03.2

Kim, S. (2015). Calling on CALL to meet individual learner needs for grammar instruction in an EFL composition class. *Multimedia-Assisted Language Learning*, 18, 97–121. https://doi.org/10.15702/mall.2015.18.3.97

Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 61, 100–140. https://doi.org/10.1111/j.1467-9922.2011.00644.x

Knoch, U., Rouhshad, A., & Storch, N. (2014). Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? *Assessing Writing*, 21, 1–17. https://doi.org/10.1016/j.asw.2014.01.001

Knoch, U., Roushad, A., Oon, S. P., & Storch, N. (2015). What happens to ESL students' writing after three years of study at an English medium university? *Journal of Second Language Writing*, 28, 38–52. https://doi.org/10.1016/j.jslw.2015.02.005

Krashen, S. (1982). *Principles and practice in second language acquisition*. New York, NY: Prentice-Hall.

Krashen, S. (1983). *Second language acquisition and second language learning*. New York, NY: Pergamon Press.

Kurzer, K. (2018). Dynamic written corrective feedback in developmental multilingual writing classes. *TESOL Quarterly*, 52, 5–33. https://doi.org/10.1002/tesq.366

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26. https://doi.org/10.1093/applin/22.1.1

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85–102. https://doi.org/10.1016/j.jslw.2009.02.001

Liao, J. (2018). The impact of face-to-face oral discussion and online text-chat on L2 Chinese writing. *Journal of Second Language Writing*, 41, 27–40. https://doi.org/10.1016/j.jslw.2018.06.005

Lynch, T. (2001). Seeing what they meant: Transcribing as a route to noticing. *ELT Journal*, 55, 124–132. https://doi.org/10.1093/elt/55.2.124

Manchón, R. M. (2017). The potential impact of multimodal composition on language learning. *Journal of Second Language Writing*, 38, 94–95. https://doi.org/10.1016/j.jslw.2017.10.008

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge: CUP. https://doi.org/10.1017/9781108333603.015

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15. https://doi.org/10.1016/j.jslw.2015.06.004

Mozaffari, S. H. (2017). Comparing student-selected and teacher-assigned pairs on collaborative writing. *Language Teaching Research*, 21, 496–516. https://doi.org/10.1177/1362168816641703

Park, J. H., & De Costa, P. (2015). Reframing graduate student writing strategies from an Activity Theory perspective. *Language and Sociocultural Theory*, 2, 25–50. https://doi.org/10.1558/lst.v2i1.24977

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21(4), 375–389. https://doi.org/10.1016/j.jslw.2012.09.004

Polio, C. (2017). Second language writing development: A research agenda. *Language Teaching*, 50, 261–275. https://doi.org/10.1017/S0261444817000015

Polio, C. & Kessler, M. (2019). Teaching L2 writing: Connecting SLA theory, research, and pedagogy. In N. Polat, P. MacIntyre, & T. Gregersen (Eds.), *Research-driven pedagogy: Implications of L2A theory and research for the teaching of language skills*. New York, NY: Routledge. https://doi.org/10.4324/9781351043281-5

Roca de Larios, J., Marín, J., & Murphy, L. (2001). A temporal analysis of formulation processes in L1 and L2 writing. *Language Learning*, 51, 497–538. https://doi.org/10.1111/0023-8333.00163

Roquet, H., & Perez-Vidal, C. (2017). Do productive skills improve in content and language integrated learning contexts? The case of writing. *Applied Linguistics*, 38, 489–511.

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29, 67–100. https://doi.org/10.1017/S0272263107070039

Schleppegrell, M., Achugar, M., & Oteíza, T. (2004). The grammar of history: enhancing content-based instruction through a functional focus on language. *TESOL Quarterly*, 38, 67–93. https://doi.org/10.2307/3588259

Serrano, R. (2011). The effect of program type and proficiency level on learners' written production. *Revista Española de Lingüística Aplicada*, 24, 211–226.

Shintani, N., Aubrey, S., & Donnellan, M. (2016). The effects of pretask and posttask metalinguistic explanations on accuracy in second language writing. *TESOL Quarterly*, 50, 945–955. https://doi.org/10.1002/tesq.323

Storch, N. (2009). The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing*, 18, 103–118. https://doi.org/10.1016/j.jslw.2009.02.003

Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies*, 10(2), 29–46. https://doi.org/10.6018/ijes/2010/2/119181

Storch, N. (2011). Collaborative writing in l2 contexts: Processes, outcomes, and future directions. *Annual Review of Applied Linguistics*, 31, 275–288. https://doi.org/10.1017/S0267190511000079

Storch, N., & Tapper, J. (2009). The impact of an EAP course on postgraduate writing. *Journal of English for Academic Purposes*, 8, 207–233. https://doi.org/10.1016/j.jeap.2009.03.001

Suzuki, W. (2017). The effect of quality of written languaging on second language learning. *Writing & Pedagogy*, 8, 461–482. https://doi.org/10.1558/wap.27291

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). Cambridge: CUP.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371–391. https://doi.org/10.1093/applin/16.3.371

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task based language learning. Insights from and for L2 writing* (pp. 217–236). Amsterdam: John Benjamins.

Torres, J. (2018). The effects of task complexity on heritage and L2 Spanish development. *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, 74, 128–152. https://doi.org/10.3138/cmlr.3770

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x

Wang, C., & Wang, M. (2014). Effect of alignment on L2 written production. *Applied Linguistics*, 36, 503–526.

Weissberg, R. (2006). *Connecting speaking & writing in second language writing instruction*. Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.6702

Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Yang, L., & Zhang, L. (2010). Exploring the role of reformulations and a model text in EFL students' writing performance. *Language Teaching Research*, 14, 464–484. https://doi.org/10.1177/1362168810375369

Yasuda, S. (2011). Genre-based tasks in foreign language writing: Developing writers' genre awareness, linguistic knowledge, and writing competence. *Journal of Second Language Writing*, 20, 111–133. https://doi.org/10.1016/j.jslw.2011.03.001

Yeh, H. C. (2014). Facilitating metacognitive processes of academic genre-based writing using an online writing system. *Computer-assisted language learning*, 28, 479–498. https://doi.org/10.1080/09588221.2014.881384

Yi, Y. (2017). Establishing multimodal literacy research in the field of L2 writing: Let's move the field forward. *Journal of Second Language Writing*, 38, 90–91. https://doi.org/10.1016/j.jslw.2017.10.010

Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51, 275–301. https://doi.org/10.1002/tesq.296

Zalbidea, J. (2017). 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *Modern Language Journal*, 101, 335–352. https://doi.org/10.1111/modl.12389

Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research*, 21, 54–75. https://doi.org/10.1177/1362168816652418

Coda

The language learning potential of L2 writing

Moving forward in theory and research

Rosa M. Manchón

University of Murcia

In this final chapter I articulate how the book contributes to advancing previous intersectional work between the fields of L2 writing and ISLA/SLA. To this end, I discuss the volume's expansion of theoretical and empirical knowledge on writing as a site for language development in terms of (a) new empirical insights provided on the learning affordances of L2 writing and the variables mediating observed effects, and (b) future research directions suggested, primarily in terms of domains to be researched, empirical questions to be addressed, and methods of inquiry to be employed.

In this final chapter I start by situating the contributions in the volume within the global framework for the study of the writing-language learning interface presented in Cumming's contribution (Chapter 2). Following from here, the main part of the chapter provides an assessment of the way in which the book collectively enhances research insights, and points to future directions capable of deepening current understandings of the language learning potential of L2 writing.

Dimensions of the connection between writing and language learning: Learning through, by, for, and with writing

It has been noted at various points in the preceding chapters that a cognitive perspective on the connection between writing and language learning has been privileged in theory and research. As a result, Byrnes (Chapter 4) reminds us, the strand "has evolved toward privileging a conceptualization of the writing-language learning interface in terms of cognitive processes and their manipulation, a decision that has affected both how it imagines and studies the act of composing itself and how it addresses learners' subsequent engagement with corrective feedback on their compositions" (p. 73). Yet, one of the defining characteristics of the present

collection is its suggested expanded range of perspectives that can and should inform work in the field. This palette of potentially relevant theoretical and analytic lenses constitute Cumming's formulation of a global framework for the study of the writing-language learning interface (Chapter 2), which he articulates as follows (p. 40. Emphasis added):

> I suggest that L2 learning while writing can be considered either narrowly or broadly, that is, as learning through, by, for, or with writing. Narrowly, L2 *learning through writing* appears to happen through cognitive problem solving and restructuring, applications and enhancements of self regulation, and collaborations with others while writing. Viewed broadly, L2 learning can be said to involve *learning by writing* (from a usage-based, activity theory, or identities perspective), *for writing* (from a perspective of motivation, purpose, or identities), and *with* writing (in complex, dynamic systems).

The theoretical chapters in Part I elaborate on the diverse global perspectives in Cumming's formulation. Thus, readers are referred to Leow's reflexive treatment of the "learning through writing" perspective (Chapter 5), to Galbraith and Al-Saadi's expanded formulation of a model of writing capable of informing future research on "learning by and with writing" (Chapter 3), and to Byrnes's carefully articulated arguments (Chapter 4) in support of a meaning-making orientation to the study of the connection between writing and language learning from a "learning by writing" perspective, and "how reconceptualizing writing as textual meaning-making might influence a research agenda at the L2 writing – language learning interface" (p. 82). What is particularly worthwhile about Byrnes's position (also echoed in other contributions to the book) is that she allows for the possibility of "common ground between a processing and a meaning orientation" (p. 76).

Importantly, the book underscores the relevance of linking the diverse perspectives in Cumming's global framework with different learning settings. For instance, Polio notes that while she agrees with Byrnes on the importance of studying writing as meaning making, "in many language learning contexts (in contrast to writing courses), writing is used specifically to facilitate language learning" (p. 383), a position Manchón and Leow (Chapter 14) also adhere to with their claim that "when the spotlight is directed at the manner in which writing in an additional language leads to language learning as part of a *language curriculum*, the inquiry ought to be situated within an instructed second language acquisition (ISLA) framework" (p. 335. Emphasis added).

Against this background, is a graphical representation of the manner in which the contributions to the volume are framed in this common research agenda and fit into Cumming's global framework. On the one hand, both theoretical/methodological chapters (Chapter 3, 4, and 16) and empirical studies (especially Chapters 12 and 13) adopt Cumming's "broad perspective" (hence shedding light
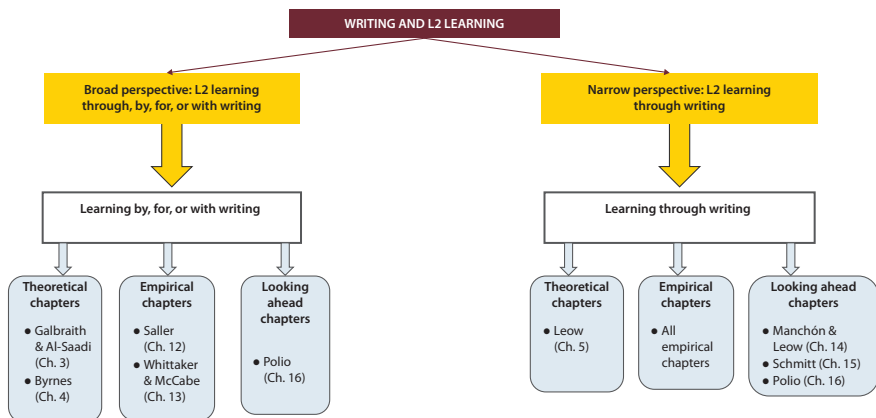
Figure 1. The chapters in the book in a global framework for the study of writing and language learning

on "learning by, for, or with writing"). On the other hand, the narrow perspective is an analytical lens present in one way or another in all empirical studies, and it is the focus of several theoretical/methodological contributions (Chapters 5, 14, 15, and 16).

In sum, the contributions to the book persuasively argue for and illustrate an expansive range of epistemological and theoretical perspectives capable of guiding the inquiry into writing as a site for language learning. Adding to this, new empirical evidence is provided on the learning affordances of L2 writing and the variables mediating observed effect, and worthy future research avenues are suggested, as I detail in the next two sections.

New empirical insights on the learning affordances of L2 writing and the variables mediating observed effects

I would like to invite readers to view this book as a concerted, collective effort to advance empirical knowledge, a global aim of the most relevance given that, as Cumming reminds us in Chapter 2, "despite the many proposed perspectives on how L2 writing may relate to and prompt L2 learning, causality remains unproven" (p. 40).

Trying to make advancements along this route, the empirical studies in Part II exemplify the diverse analytic lenses that can be applied to the study of writing as a site for language learning, illuminate the myriad of individual, educational, and task-related variables that (may) mediate short-term and, importantly, long-term language learning outcomes of diverse forms of writing (including writing tasks performed in controlled experimental settings, as well as writing assignments that

were part of the participants' educational experience), in diverse environments (including pen-and-paper and digital writing), performance conditions (writing individually and/or collaboratively), and instructional settings (academic settings – including secondary school and college level institutions- as well as out-of-school contexts).

Key empirical contributions of the book are synthesized next. Reference will also be made to how this new evidence connects with extant theoretical predictions on how L2 writing may advance L2 competences.

The language learning potential of writing from the perspective of the affordances of writing vs. speaking

The language learning affordances of writing are in great part predicated on the time-nature of written output: The greater availability of time in the writing condition, as opposed to the stricter time limitations associated with the real-time production of oral language, is thought to be more propitious for language learning in terms of facilitating the control and allocation of attentional resources, and the resulting enhanced attention to language-related concerns (see Manchón & Williams, 2016; Williams, 2012). This greater attention to language is precisely the purported driving force leading to language learning outcomes.

One way of testing this prediction is by comparing task performance across modalities. Hence, building on the extant body of task-modality studies (recently reviewed by Manchón & Vasylets, 2019), several chapters in the book (Chapters 6, 7, 8, and 9) represent advancements in the modality-sensitive research agenda (originally propounded by Harklau, 2002. See Chapter 1) with their focus on the study of task complexity and task repetition across modalities. Four sets of empirical findings in the research reported in these chapters are worth highlighting:

1. One common insight (see especially Chapters 8 and 9) is that modality plays a more decisive role than task complexity and it constitutes a more propitious environment for bringing about task complexity effects on production and language learning (see Vasylets et al., Chapter 8). For instance, Zalbidea (Chapter 9) found that, although the two task modality conditions in her study resulted in substantial noticing and incorporation of the target form, the writing condition led to more accurate form incorporation into the participants' output, a finding that the author explains with reference to current theoretical predictions. Thus, Zalbidea claims that "the slower-paced, visual and more permanent nature of the written modality provided greater opportunities for participants to identify mismatches between their output and the model feedback" (p. 224), adding that "This lower-pressure environment […] may have

allowed learners more time and support to successfully disentangle the complex form-to-meaning relationships" (pp. 224–225) of the target, which resulted in the more accurate use observed.

2. The empirical research reported in the book additionally provides empirical validation for the theoretically predicted modality-dependency of task-repetition effects (cf. Manchón, 2014). Thus, in their study of task repetition across modalities, Sánchez, Manchón, and Gilabert (Chapter 6) found task repetition to show its own idiosyncrasy in writing, a finding further reinforced by Amelohina, Nicolás-Conesa, and Manchón (Chapter 7) in their longitudinal study of task repetition in writing with the availability of feedback. In line with Zalbidea's findings reported above, Sánchez et al. interpreted their data as suggesting that writing constitutes a more favorable environment than speaking for increasing attention to language, which in their study elicited the use of more complex language in some dimensions of syntactic and lexical complexity. In line with Zalbidea's arguments, the authors attributed this finding to the differential temporal nature of oral and written communication in the terms specified above.

     It should be mentioned in passing that these two task repetition studies (reported in Chapters 6 and 7) also shed light on concerns that have been at the heart of the research interests in the task repetition literature. Thus, Sanchez et al.'s study provides insights in support of the predicted difficulty in anticipating task repetition effects, and sheds new light on the divergent predictions about the proficiency-dependency of task repetition effects (see Bygate, 2018). Similarly, Amelohina et al.'s investigation of the appropriation of written corrective feedback (WCF) over time contributes novel insights given that, as the authors discuss in the framing of their study, feedback also constitutes part of professional discussions on task repetition (see Bygate, 2018; Ellis, 2009).

3. A third relevant message from the task-related studies in the book is the reiterated conjecture that different language modalities may bring about diverse language learning benefits. For instance, Zalbidea interpreted her findings as evidence of differential affordances of speaking and writing for grammar development, and Vasylets et al. (Chapter 8) interpreted theirs as pointing to a differential learning potential for lexical development across modalities, a claim that can be linked to Sánchez et al's findings of greater attention to lexis in the writing mode (although this effect was observed only for the high proficiency writers in their study).

Collectively, these insights and tenets open up avenues worth exploring in future research agendas, as further discussed by Schmitt (Chapter 15) and Polio (Chapter 16) in their respective contributions (see below).

Learning through writing: New insights on the processing dimension
of L2 writing and associated learning effects

The essence of the psycholinguistic rationale for the language learning potential of L2
writing is primarily cognitive in nature, as Cumming reminds us in Chapter 2 (p. 37):

> When writing, if lacking appropriate words or language forms in an L2 or L1,
> people can pause to regulate their writing performance, search their memories
> for relevant resources, apply heuristic search strategies across first or second lan-
> guages, and restructure tentative verbal formulations to identify and resolve per-
> ceived problems and so confirm or extend their knowledge […]. Such episodes of
> self-control, reflection, and cognitive restructuring provide potential opportunities
> for learning the L2 while writing it.

In line with these predictions, three studies in the book (reported in Chapters 9, 10,
and 11) advance current understandings of why and how the processing dimension
of writing (in individual – Chapters 9 and 10 – and collaborative – Chapter 11 –
writing conditions) connects to and leads to language learning. They do so by
providing novel research insights and by recommending and illustrating method-
ological innovations.

From the perspective of research methodology, it is worth mentioning López
Serrano, Roca de Larios, and Manchón's (Chapter 10) theoretically grounded an-
alytical framework for the analysis of linguistic processing while writing (see also
López Serrano, Roca de Larios, & Manchón, 2019). It is pertinent to note that
their framework is partly based on Leow's (2015) model of instructed SLA, which
Leow himself connects to the analysis of writing and language learning in his own
contribution to the book (Chapter 2). Also relevant from a research methodology
perspective, Stiefenhöfer and Michel (Chapter 11) offer an illuminating account
of methodological considerations in the study of collaborative writing process in
digital environments, including the affordances of data triangulation (also present
in Zalbidea's study of writing processes in individual writing conditions) and the
problems experienced and solutions adopted in their own research endeavor.

From the perspective of new research evidence, the main research insights
provided in the process-oriented studies included in the volume can be summa-
rized as follows:

1. These studies provide convincing evidence for the theoretical predictions on
   how writing promotes noticing and cognitive comparison processes. For in-
   stance, Zalbidea (Chapter 9) interpreted noticing processes in her data as in-
   dicative of effects for restructuring in the writing modality, where more learner
   generated focus on form was observed. Similarly, in their study of collaborative

writing processes, Stiefenhöfer and Michel (Chapter 11) suggest that the saliency and permanence of digital writing may lead to language learning through heightened attention to form and the process of cognitive comparison, adding that the "permanence of output (written composition and text chat conversation) may serve as a source for vocabulary and grammar, available for cognitive comparison and subsequent writing" (p. 259).

2. Closely linked to the previous point is the evidence obtained on the connection between levels of depth of processing (DoP) and language learning in individual (López Serrano et al., Chapter 10) and collaborative (Saller, Chapter 12) writing conditions. This connection is another important concern in the theorizing on writing as a site for language learning. Thus, the participants in Saller's study engaged in complex meaning negotiations and deep problem-solving behavior in their collaborative writing activity. Similarly, López Serrano et al.'s study illuminates the proficiency-dependency of the levels of linguistic processing while writing individually. Even more consequential, the authors uncover the complexities involved in understanding this linguistic processing and potential effects on learning, which I interpret as a challenging invitation for a critical reconsideration of the (mostly untested) assumed equation between deeper levels of processing and higher levels of learning. More precisely, López Serrano et al. explain how their pre-intermediate and intermediate learners engaged in deep output processing mostly for compensatory purposes (i.e. to compensate for linguistics deficits), which the researchers claim resulted in noticing holes and gaps, hypothesis formulation via activation of L1 and L2 knowledge, and self-evaluation (via metalinguistic refection) of their own output. Yet, López Serrano et al. add that for these writers, because of the limited prior knowledge they could resort to, the "chances of consolidating partially acquired meaning-form relationships or creating new ones were reduced" (p. 247). As a result, they argue, "compensatory episodes would support the claim that *while deeper levels of processing may increase the likelihood of achieving higher levels of awareness as understanding, this is not always the case* (p. 248. Emphasis added).

These observations distinctively point to the relevance of expanding process-oriented L2 writing research before firmer conclusions on the purported connection between linguistic processing while writing and language learning can be safely reached. It should be recalled in passing that such connection has already been problematized on the basis of the findings obtained in studies looking into the correlation between depth of processing of feedback and accuracy improvement in revised texts (e.g. Cerezo, Manchón, & Nicolás-Conesa, 2019).

3. The process-oriented studies reported in Part II shed additional light on the learning that accrues in writing and do so in a dual way as, on the one hand, these investigations put theoretical tenets to the empirical test, and, on the other, they advance new tenets to be tested in studies of individual and collaborative writing.

Regarding individual writing, the data obtained by López Serrano et al. (Chapter 10) led them to suggest that "while writing challenging texts, L2 writers might not only be consolidating their L2 knowledge but likely also establishing new, still underdeveloped form-meaning spaces for future development" (p. 249), which is a finding also corroborated in Zalbidea's and Stiefenhöfer and Michel's studies. Their reflections also connect with some of the arguments put forward by Leow in Chapter 2 regarding writing, feedback processing, and language learning. López Serrano et al. suggest (p. 248):

> the learning potential can also be found in the traces that the cognitive effort involved in the activation of L2 knowledge (especially at higher levels of processing) may leave in learners' memory. If these traces remain or are activated by the time feedback is provided, (see Leow, Chapter 5, this volume), learners may greatly benefit from either testing their hypotheses about the L2 or from receiving correct L2 forms in the written corrective feedback provided.

López Serrano et al. further link their conjecture to the relationship between input and output processes in SLA (Gass, 2010), which coincidentally is a key finding in Whittaker and McCabe's study reported in Chapter 13, to be discussed in a later section. For now, let me simply note the relevance of Whittaker and McCabe's observations about depth of processing given the longitudinal and classroom-based nature of their research, which represents a clear departure from the controlled, laboratory-type studies that constitute the bulk of extant research on writing processes. On the basis of their data, Whittaker and McCabe argue that the unique characteristics of writing purported to be potentially conducive to language learning can be well exploited in instructional settings in which L2 users learn content through an additional language, as was the CLIL context they investigated. In their view, these learning contexts "are a site for cognitively demanding writing tasks, in which students can engage as part of the learning of a subject" (p. 327). Accordingly, they conclude that "incorporating writing systematically into these classes would lead students to discover gaps in their language knowledge, and explore and expand their resources, with this process of reflection and rewriting leading to "deeper learning" '"(p. 327). These are very relevant suggestions for future work on writing processes and their language learning potential, and readers are referred to the arguments in Chapter 5 (Leow) and Chapter 14 (Manchón & Leow) on the relevance

and implications of adopting an instructed second language acquisition (ISLA) perspective in the quest for answers on the connection between writing and language learning (see also below for suggestions for future research by Schmitt and Polio).

Further insights on the learning that accrues in writing comes from the two studies on collaborative writing, especially regarding their observations on the role of interaction in bringing about language learning. According to Stiefenhöfer and Michel (Chapter 11), collaborative writing supports L2 learning through both the potential of writing itself in the terms discussed so far, as well as from the benefits that derive from the kind of interaction that characterizes collaborative writing. In their study, this potential manifested in the range of interactional patterns observed and the participants' "engagement with the writing task and the writing process, often simultaneously through observing, editing, or discussion with the partner" (p. 275). Saller (Chapter 12) also interprets his evidence of dyads deliberating upon complex meaning-making decisions from the perspective of the language learning affordances of such interaction. He adds relevant observations for future work, including the scarcity of research on the nature of dyadic metatalk in what he calls "complex meaning-making environments", the limited research conducted with advanced writers, the absence of a longitudinal perspective in extant research, and the relevance of inspecting learners' perceptions of collaborative writing since, in his view, "these perceptions strongly influence the effectiveness of collaborative writing tasks" (p. 284) and, hence, any learning that may derive.

Learning by writing: Language learning and the meaning-making activity that is criterial to writing

In an earlier section I referred to Byrnes's (Chapter 4) enforced arguments in support of a meaning-making orientation to the study of the connection between writing and language learning. In essence, Byrnes contends that this research agenda would entail a search for questions regarding "how L2 writers go about settling on the details of their language choices, what those choices are with regard to their meaning-making function in a textual environment, and how they change and develop over the course of the writers' evolving textual abilities" (p. 81). Readers can see the application of this research approach in the empirical studies reported in Chapters 12 (Saller) and 13 (Whittaker & McCabe). Thus, it is worth reiterating Saller's finding regarding the meaning and content negotiations privileged in the kind of complex composition tasks his advanced L2 users completed collaboratively. Adding to this, the longitudinal, classroom-based study conducted by Whittaker and McCabe (Chapter 13) in an ecologically valid environment (as was also the case with Saller's and Amelohina et al.'s studies) looked into the role of

writing in bringing about language development over time. Coming back to the input-output connection referred to in an earlier section, their data distinctively show that the CLIL students investigated learned to use nominalization through input processing (reading history texts) and output practice (writing over 4 years), which is precisely the experimental condition (input/output cycles) in Zalbidea's (Chapter 9) controlled study. Whittaker and McCabe's research illustrates Byrnes's suggestion to investigate the "linking of reading and writing in order to foster literacy development" (p. 85), the "integration of writing into content-based approaches to language teaching" (p. 85), the emphasis on how source-based writing "inherently involves reading" (p. 85), and the relevance of opting for "composing assignments and/or tasks that require the use of textual sources" (p. 85). Illustrating these research directions in a real classroom over time, Whittaker and McCabe observed notable developments in their participants' texts over the four years of the duration of their study, a key finding that the researchers attributed to the distinctive meaning-making opportunities that are possible in the environment of writing. In their view, such learning opportunities "can only develop through *interaction with written texts, both through their consumption and, especially, through their production*" (p. 323. Emphasis added). Importantly, in the context investigated there was no overt writing instruction, or teaching of the linguistic targets, hence their speculations that "studying history in an FL provided strong input for writing development, with several characteristics of the discourse of history calling for its use" (p. 324) or that "In studying history through English, the students have been exposed to its nominalizations and technical abstractions, which they begin to incorporate into their own texts" (p. 325).

I consider these to be novel and truly relevant findings that open up new research avenues on the role of intentional/incidental learning through writing in classroom settings, on the role of reading in bringing about learning through writing, and on the relevance of adopting the longitudinal perspective in classroom-based research, as discussed below.

New directions for future research agendas

In this section I synthesize the suggestions put forward in the theory and research reported in the book regarding new areas in need of further scrutiny. Critical methodological considerations for the suggested future research agendas will be accounted for in the following section.

Future inquiry into writing processes and effects on learning

The process-oriented empirical studies in Part II, together with some of the position pieces in Part I and Part III, recommend future inquiry on the processing dimension of writing. I shall not reiterate here the future avenues signaled in the empirical studies reported in Part I that I have already referred to in the previous section. Instead, the contributions in all chapters are discussed in relation to the framework and guidelines proposed in Manchón and Leow's chapter (Chapter 14) as well as with specific reference to the suggestions for further research proposed in Schmitt's and Polio's chapters (Chapters 15 and 16, respectively).

Manchón and Leow's comprehensive, ISLA-oriented agenda on writing processes point to two main future developments. First, they argue for an expansion of writing conditions, including more traditional and emergent individual and collaborative writing in diverse environments, the latter crucially including pen-and-paper and digital writing in recognition of the fact that writing has gradually become more digital and this has impacted writing processes and products in fundamental ways (see Hafner, Chik & Jones, 2015; McKee & DeVoss, 2007; Oskoz & Elola, 2020). Second, Manchón and Leow put special emphasis on the time-distributed nature of writing and the resulting implications for the study of writing processes as they unfold over time, a research endeavor that, the researchers argue, entails incorporating innovations in methodological procedures and acknowledging crucial ethical dilemmas to be solved. Along similar lines, Polio (Chapter 16) argues in favor of "methods that can capture the longer process of real life writing tasks" (p. 391), adding that "Outside of testing contexts, when time pressure is less and when there is access to sources, writers likely draw on more explicit grammatical knowledge, editing strategies, dictionaries, grammar checks, corpora, and readings. If we are to suggest that writing allows for a focus on and reflection about form, we need to understand these real life processes" (p. 391).

Recent studies by Hort (2017) and Seror (2013) represent worthy empirical attempts along the directions pointed in Manchón and Leow's and Polio's chapters. Thus, Hort (2017), asked her participants to keep a "digital diary" on their writing everyday with the aid of an application that allowed them to incorporate text and photos. The participants were instructed to write (*inter alia*, about where they were, how they felt, and what they did) every time they engaged in an action related to a given assignment (2-week writing assignments and a final writing task written over 10 weeks). Similarly, Seror (2013) used screen capture technology (Screencast-O-Matic) to analyze the time-distributed nature of the writing process in an attempt to "unobtrusively gather, store and replay what have traditionally remained hidden sequences of events at the heart of L2 writers' text production" (p. 1). The

participants in the study were invited to record themselves at their own discretion while completing primarily writing assignments for their writing class, which allowed the researcher to collect data on the participants' process while completing full short class writing assignments as well as while writing parts of longer texts. The researcher could thus shed light on what he termed "the sequential nature of specific composition processes and strategies" (p. 7)

In short, future work on writing processes ought to expand both its scope – so that digital writing becomes more prominent in research agendas – and the time-frame conditions under the spotlight – in an attempt to account for the sequential, time-distributed nature of writing events.

Future inquiry into learning that accrues in writing: The role of tasks

Understanding more fully the learning that accrues in writing is another fruitful line of research to pursue reiterated in several contributions to the book. It should be recalled that the theoretical predictions on how and why writing may lead to language learning (e.g. Manchón & Williams, 2016; Williams, 2012) allow for a range of potential outcomes, as Cumming synthesizes in Chapter 2 (p. 32): "The permanency, self-controlled pace, and expectations for precision of expression in writing can prompt learners to attend explicitly to language forms and meanings together to refine and consolidate their L2 knowledge and skills".

Cumming also mentions the learning that may derive from having access to model texts and peer feedback. As for the latter, readers should recall the reference above to the learning potential of the interaction that is criterial to collaborative writing (see contributions by Stiefenhöfer & Michel, Chapter 11, and Saller, Chapter 12). As for the role of model texts, readers are referred to the earlier discussion of the reading-writing connection and associated learning outcomes in Whittaker and McCabe's study (Chapter 13). Finally, Cumming mentions the possibility of restructuring "complex, dynamic, interacting language, discourse, and semiotic systems through repeated usage and personal agency in communicative interactions" (p. 35). This is a point Saller (Chapter 12) elaborates at length in his discussion of the outcomes of the deliberations his participants engaged in while completing writing tasks collaboratively, arguing that these deliberations may push advanced L2 learners to restructure their syntactic output during text generation in order to meet the communicative demands of challenging writing tasks, which Byrnes (2014) has referred to as "the ability to make situated linguistic choices" (p. 87).

In addition to the empirical evidence provided, several chapters in the book offer a wealth of suggestions for moving research agendas forward, specifically in terms of vocabulary and grammar learning. Regarding vocabulary, and in attempt to "bring the concerns of vocabulary research and writing research into closer

alignment" (p. 357), Schmitt advocates a stronger word level focus in future inquiry in the domain. From a writer-internal, process-oriented perspective, and being fully cognizant of the crucial role of lexical searches in the process of text creation (see Murphy & Roca, 2010), Schmitt notes the need to "resolve what is essentially a chicken or the egg dilemma – does the process of solving lexical problems generate new knowledge about specific words which learners can then add to their lexical entries or does success with problem solving depend on the existing level of detail in writers' lexical entries?" (p. 371). One possible answer, derived from López Serrano et al.'s data (Chapter 10), is that it might be a question of the interaction between the L2 writer's proficiency level and the cognitive demands of the task at hand. Thus, on the basis of their data, López Serrano et al. hypothesize that the more expanded linguistic resources their advanced participants could draw on facilitated their deeper linguistic processing while attempting to find a match between their intended meaning and the linguistic means needed to express their ideas in the L2. In contrast, they argue, their lower-level learners, "whose L2 knowledge or access to it is more limited and who may suffer from cognitive overload, argumentative tasks may not be the most appropriate tasks to take full advantage of the language learning potential of writing" (p. 247). In this respect, it is important to mention Byrnes's emphasis on the developmental nature of writing capacities and, hence, on the relevance of "the kind of textual meaning-making that L2 writers are asked to engage in *at different points of their long journey to becoming competent writers*. In other words, the *increasingly sophisticated awareness these writers have regarding the functional – that is, the communicative – consequences of deploying certain lexicogrammatical resources of their L2* must be of central concern to researchers." (p. 78. Emphasis added).

From a more learner-external perspective, a reiterated, loud claim in the book is the imperative to understand more fully the role of tasks in bringing about language learning through, by, and with writing. For instance, Stiefenhöfer and Michel (Chapter 11) contend that in order to take full advantage of the learning potential of digital, collaborative writing, research has to establish "how tasks should be designed, how pairs should be formed, and what communication modes should be encouraged in order to elicit fruitful interaction that potentially benefits SLA." (p. 273). Polio (Chapter 16) equally draws attention to the relevance of considering writers' engagement with the task and of expanding "our repertoire of writing contexts, including the prompts or activities, to those that hold more investment for research participant" (p. 390). Again in terms of vocabulary learning through writing, and taking a decidedly curricular, classroom- based approach (as also done by Leow, Manchón, & Leow, and Byrnes in their respective contributions), Schmitt categorically argues that the first requirement for vocabulary learning through writing is that words need to be produced, which means that tasks have to be designed in

such a way that they motivate L2 writers to use new or partially known vocabulary (this addresses the "size challenge" discussed at length in Schmitt's contribution), and to do so repeatedly, in a range of contexts and across modalities. Schmitt argues that only then will it be possible for researchers to "investigate the contributions of different approaches to task design, genres, and feedback to the incremental development of individual words and to the wider mental lexicon" (p. 377). She adds very relevant suggestions for task repetition studies, where researchers are invited to pay greater attention to "both the quantity and quality of the encounters that students have with individual words" (p. 375).

The book offers readers numerous additional proposals regarding central task-related concerns for future research agendas. I shall focus on just three very relevant ones in Polio's analysis of future work on grammar learning through writing (Chapter 16). The first concerns the study of multimodal writing, an area of research in which the key question would be to ascertain the extent to which "different types of multimodal writing affords a focus by on language, in general, and grammar in particular" (p. 390). In order to do so, Polio recommends research approaches in which introspection techniques (such as think-aloud protocols or stimulated recalls) are used while participants complete controlled multimodal tasks as this "would allow us to evaluate the concern that multimodal tasks might not facilitate language learning" (p. 390. See Manchón, 2017, for reservations about establishing a direct link between multimodal writing and language learning).

A second task-related research direction in Polio's analysis relates specifically to task complexity studies and the question as to whether writing tasks can be manipulated to promote grammatical development. The question Polio raises is whether the more complex language that results from given tasks and genres "is due to some type of cognitive requirement or an affordance of the writing prompt (i.e., the task or genre requires more complex language to communicate)" (p. 393). She further notes the relevance of observing the "long-term effects of […] various tasks and genres on the development of grammatical complexity" (p. 393), as well as the need to tackle another research dilemma: "If we are to argue writing promotes development, we need to show that a task that elicits more complex language leads to learning" (p. 393).

A third line of research concerns Polio's suggested areas of interest in future task-modality studies. Her suggestions can be linked to Byrnes's (Chapter 4) call to answer questions on what she refers to as "separate and/or integrated development of writing within all modalities" (p. 84). In this respect, Polio notes the scarcity of research that track learners' development in oral and written language simultaneously, an important gap in her view in recognition of two relevant empirical questions in the domain. One is whether written grammatical development precede oral grammatical development, a key question that Polio claims needs to be answered in order to maintain that writing facilitates language learning. The second empirical

question is whether writing activities promote oral grammatical competence and hence whether what is learned in writing can be transferred to speaking. Polio considers this to be most relevant question to be posed.

The role of individual differences

The role of individual differences is a key dimension of the writing-language learning interface also requiring additional empirical scrutiny. According to Cumming (Chapter 2), issues that warrant further attention include the following (p. 41):

> whether the potential for L2 learning through writing is qualitatively different at different levels of L2 proficiency or even whether it might be quantitatively different in the sense that some kinds of learning […] require a certain level of L2 writing fluency and foundation lexical resources, as Cummins' (1984) "threshold hypothesis" proposed. Similarly, are certain ages of maturity, levels of education, or literate abilities required to facilitate L2 learning through L2 writing? Further, might there be a point at high levels of L2 proficiency when such processes exhaust their usefulness?

The empirical chapters in the book distinctively point to a clear association between higher levels of proficiency and greater potential learning benefits, as manifested, for instance, in the participants' use of language (especially regarding use of lexis. See Chapter 6) and their depth of processing (see Chapter 10). Additionally, Saller (Chapter 12) attributed the meaning and content negotiations in his data (and resulting learning effects) to the participants' high proficiency level. These are, nevertheless, tendencies that require further empirical validation. In doing so, future work should be mindful of the relevance of investigating all proficiency levels, because, as Byrnes (Chapter 4) rightly reminds us, failing to do so would severely limit the scope and implications of work in the domain. In her own words (p. 75):

> Not only are even beginning-level writers at all educational levels quite capable of considering texts as a mode of communication that differs from oral language use and of adjusting their language resource use accordingly; more consequential, not allowing for that capacity early on seriously undermines the entire argument for an L2-writing-language learning link with obvious consequences for investigating it substantively.

Along these lines, Manchón and Leow (Chapter 14. See also Leow & Manchón, forthcoming) argue for an expansion of the populations investigated in recognition of the preponderance of college students (often with a background in linguistic or language studies, which severely limits the generalizability of finding, as the authors themselves explain) or teenager participants in the extant research, while studies

with younger populations are underrepresented in the field (see Michel, Kormos, Brunfaut, & Ratajczak, 2019 for a recent exception).

One further item for future work on individual differences is suggested by Cumming (Chapter 2): Variations in cross-linguistic combinations, whose study could be approached "either through within-subjects designs involving the same people composing similar tasks in first and second languages, or through comparisons between L1s and L2s that are either highly similar or different in their scripts and discourse structures" (p. 41).

Finally, the two studies in the volume on collaborative writing underscore the relevance of investigating additional affective and attitudinal individual learner factors. Thus, Saller (Chapter 12) discusses the role of intersubjectivity and the participants' own perceptions of the attainability of tasks as relevant variables in explaining any learning that may derive from the engagement with writing tasks. Similarly, Stiefenhöfer and Michel (Chapter 11) draw our attention to the likely inhibitory effect that the online environment could have "for some students, who might feel exposed while writing. Those that do not feel comfortable with the idea of having somebody look over their shoulder as they write might deliberately choose the role of editor, leading to less participation" (p. 272).

The role of instruction in promoting vocabulary and grammar learning

Schmitt and Polio's contributions offer a wealth of directions for future studies that target how specific interventions promote vocabulary and grammatical development.

Regarding vocabulary learning through writing, and in line with the above mentioned word level focus for future research that Schmitt (Chapter 15) advocates, many worthy avenues worth pursuing are formulated, of which I shall highlight two. One is to "investigate whether explicit instruction that calls attention to various aspects of word knowledge, and provides opportunities for L2 writers to build more elaborate lexical entries […] can lead to better quality evaluation of vocabulary choices during lexical retrieval episodes" (p. 371). Another worthy avenue to pursue, and one that connects with the task repetition studies in the book, is Schmitt's argument in favor of tasks that approximate exact repetition designs as this would likely facilitate the use of the same words multiple times.

Regarding grammar, Polio underscores the relevance of ascertaining whether writing instruction can be linked to grammatical development given the open questions that exist on the link between specific instructional techniques and grammatical development. She also poses a question that I consider of the most relevance: To ascertain whether or not asking language students to write, providing feedback on

their writing, and asking them to rewrite/revise their texts on the basis of the WCF received is the most effective, less time-consuming, and more ecologically-valid way of fostering L2 grammar learning through writing. Problematizing this approach, Polio lists other types of potentially useful pedagogical interventions, including pre- and post-task metalinguistic instruction, provision of model texts, or intervention that focus on the grammatical features of authentic readings. As for the latter, readers are referred once again to the input/output, reading/writing connections discussed above in relation to Whittaker and McCabe's study.

Future empirical research agendas: Research methodology considerations

Readers can also find a wealth of research methodology considerations for future work in most of the contributions to the volume. I have chosen to comment on those that I consider to be specially relevant in advancing empirical work in the domain.

Refinements of constructs and analytical approaches

As a general observation, Cumming (Chapter 2) argues that "methodological and conceptual refinements […] might lead researchers to produce evidence about the benefits of L2 writing for L2 learning" (p. 41), a challenge taken up in various contributions. Thus, Schmitt (Chapter 15) calls for refinements in research tools and analytic procedures in diverse areas of research related to writing and vocabulary learning. For instance, in studies on the impact of task repetition on the development of word knowledge, she advocates "finer grained analyses than offered by commonly used error classification and lexical complexity measures" (p. 374), a mandate of the most relevance given the prominence of errors and CAF dimensions as outcome measures in many areas of research on how writing may lead to language learning. Schmitt's detailed critical observations of extant classifications and analytic approaches, together with her well-grounded suggestions for overcoming current problems, are therefore especially relevant. Readers are also referred to Leow's (Chapter 5) recommended refinements in the analysis of errors, especially his suggestion "to avoid using in future studies global scores (e.g., number of errors per 100 words, number of error-free T-units, etc.) and analyze L2 writers' actual errors produced on compositions […] submitted during the semester" (p. 111).

Manchón and Leow (Chapter 14) also take up Cumming's global challenge for needed methodological refinements. Out of the numerous considerations that readers can find in their chapter, I would like to highlight their call to problematize our constructs when new populations of instructed learners are investigated,

a suggestion that is the direct consequence of their claimed expansion of the range of populations to be studied (see above). They also invite researchers to be mindful of the available evidence on the complexity that characterizes the processing, problem-solving activity that is criterial to writing, a complexity that, the researchers argue, must be fully acknowledged and be integrated in the analytic tools and coding schemes guiding future empirical inquiry. Two examples in the book testify to this. In terms of collaborative writing, Stiefenhöher and Michel (Chapter 11) call for "a more dynamic approach to classifying pair interaction, as participants showed changing patterns throughout the tasks" (p. 275). Regarding individual writing, López Serrano et al. (Chapter 10) call for needed refinements in the analysis of depth of processing (DoP), a construct that has become central in the domain. They urge for more complex approaches on account of the fact that "the potential role DoP may play in promoting language learning via free writing tasks is not a black and white issue, but one that can be better understood if it is contemplated in connection with the strategic orientation (compensatory or upgrading) adopted by the writers when addressing their self-generated problems" (p. 249).

Relevance of data triangulation and the use of mixed-methods research approaches

Further reiterated global conclusions relate to the benefits to be obtained from data triangulation and the use of mixed-methods research approaches.

From the perspective of "learning through writing", Leow (Chapter 5) argues that "researchers need to address the dearth of concurrent data on the cognitive processes employed by L2 writers at all levels of proficiency during both the composing and revision phases of the writing process" (p. 110) as this "will allow researchers to avoid making assumptions on how L2 writers process the L2 data or WCF and clearly lead to a better understanding of the writing-to-learn process" (p. 111). Yet, Cumming (Chapter 2) suggests going beyond this kind of analyses of metacognition via self-reports so that research can "address more basic, implicit or tacit aspects of cognition such as attention, neural processing, or self-regulation through, for example, brain imaging, eye-tracking, or response-time methods". In fact, several chapters (see especially contributions by Zalbidea, López-Serrano et al., Stiefenhöfer & Michel, and Manchón & Leow), support this expansion of methodological approaches as well as the benefits of instrument triangulation. As an example, Stiefenhöfer and Michel (Chapter 11) conclude from the study that data triangulation procedures allow more subtle analysis, as was their case where combining eye-tracking with stimulated recall interviews proved to be particularly valuable in the identification of noticing processes.

Finally, several voices in the book recommend the adoption of qualitative approaches when investigating the writing-language learning interface. For instance, Polio (Chapter 16) encourages an expansion of the scope of inquiry into grammar learning through writing using new methods. She highlights the advantages that could derive from detailed case studies, which, in her view, might provide richer insights into whether or not given structures were learned. She also makes recommendations in terms of data triangulation and the use of mixed methods research as "an ideal way to better understand how language does or does not develop in writing classes because we can link previous quantitative measures with data from classroom observations, interviews, and artifacts" (p. 391). Similarly, although from a different perspective, Cumming (Chapter 2) wonders whether "distinctly qualitative methods of inquiry, such as life-histories or ethnographies of classrooms or multilingual workplaces" could "illuminate L1/L2 interactions involving significant learning at key incidents in the lifespan and within specific discourse communities" (p. 41).

Adopting an educational, curricular approach: The needed longitudinal perspective

The book makes an appealing invitation for future researchers to go into real classrooms and hence to adopt educational and curricular perspectives in their research endeavor. As further detailed below, adopting this research orientation would entail opting for longitudinal studies.

The educational perspective is convincingly articulated by Byrnes with her argument that "evidence from finished compositions cannot offer proof that writing caused language learning to that level" (p. 83). Instead, she argues, "if such literate language use is consistently attained in an L2 educational setting that deliberately fosters it within its entire program, then it should yield insights regarding conditions that are favorable to the hypothesized relationship" (p. 83). She goes on to discuss the methodological challenge that adopting such perspective would entail, which she sees as "that of reconciling the complexly dynamic nature of language use and development – its complex systems nature – with the unavoidably restricting exigencies of educational settings" (p. 90).

Importantly, Byrnes claims that the construct of curriculum is "one way of recognizing these countervailing dynamics" (pp. 90–91), a position that is also at the basis of the curricular orientation for future work advocated by Leow (Chapter 5) and Manchón and Leow (Chapter 14). For instance, Leow argues that situating future work on writing-learning interfaces within a curricular approach "is of paramount importance if researchers would like to extrapolate their empirical findings to the instructed setting" (p. 98). Manchón and Leow reinforce the same idea with their

claim that future studies aiming to provide pedagogical implications must situate their designs within the language curriculum in specific classroom settings, which entails moving from one-shot designs and instead design longer-term investigations in order "to simulate the natural syllabus of a language class" (p. 351). In their view, following this path would allow to "capture language learning through curricular, repeated, and extended opportunities to write, and address writing processes while performing writing activities that spread over different time spans" (p. 351).

The longitudinal perspective is also recommended for studies of task repetition in order to ascertain the long-term effects of the procedure (see general suggestions in Sánchez et al.'s chapter, and specific observations about vocabulary learning through task repetition in Schmitt's chapter). Importantly, Amelohina et al.'s study (Chapter 7) of task repetition in the writing mode with the availability of written corrective feedback (WCF) shows the relevance of zooming into the potential differential appropriation of WCF over time (seldom investigated thus far), which the longitudinal design of their research allowed them to capture. The study further reinforces the relevance of the educational perspective mentioned above, in their case studying WCF appropriation in ecologically valid contexts in which the teacher's provision and students' engagement with WCF was an integral part of the participants' learning experience.

Polio (Chapter 16) also embraces a longitudinal approach in the study of grammar learning through writing as the field still has to ascertain the long-term facilitative effects of writing for grammar learning. Similarly, Byrnes (Chapter 4) discusses the benefits to be obtained from a longitudinal inquiry into the L2 writing-language learning interface, especially to trace the developmental changes in the L2 writer's meaning-making capacities. This approach is ultimately linked to the educational perspective that she advocates, as discussed above, and thus to the empirical question of "how we are to imagine development in L2 writing and L2 language use in its extended evolution, under what configurations of educational practices for what learners with what realistic learning outcomes for what languages" (p. 86).

Concluding remarks

This is a book that celebrates the achievements of a rather recent but nevertheless vibrant area of scholarly work at the intersection between SLA and L2 writing studies, while at the same time it attempts to push this domain of research theoretically and empirically. Theoretically, the main contribution the book wishes to make is the expansion of the theoretical perspectives capable of informing scholarly work in the domain: The contributions to the book persuasively argue for and illustrate an expansive range of epistemological and theoretical perspectives capable of guiding

the inquiry into writing as a site for language learning. Empirically, the book advances current professional discussions with a wealth of new empirical insights that serve to both support and problematize theoretical positions and predictions on how and why writing may be conducive to language learning. The book compiles empirical studies that are framed in a common research agenda, they exemplify the diverse theoretical lenses through which we can inspect writing the language learning potential of L2 writing and, collectively, these studies shed light on the many variables at work in bringing about short- and long-term language learning by, with, and through (individual and collaborative) writing, in pen-and-paper and digital environments, both in experimental conditions and when L2 users engage in writing task completion over time as part of their educational experience. Empirically, the book additionally signals fruitful avenues to be explored by pointing to needed methodological and conceptual refinements, underscoring pending empirical questions to be addressed, and problematizing and expanding methodological approaches to be employed.

In sum, the present volume represents a collective attempt to expand theoretical and empirical knowledge on writing as a site for language learning, while at the same time it calls for a critical reflection of where future professional initiatives should be directed. This critical positioning in part entails what Byrnes (Chapter 4) refers to as a collective, professional critical consideration of "the extent to which lacunae in our knowledge regarding the writing-language learning interface might reflect conceptual and empirical habits of mind that by now merit careful re-evaluation" (p. 74).

Ultimately, it is hoped that the volume, through the double lens of enhancing research insights and critically pointing to needed developments, constitutes a solid springboard for future professional initiatives aimed at deepening current understandings of how and why writing in an additional language can serve to advance language competences.

References

Byrnes, H. (2014). Theorizing language development at the intersection of "task" and L2 writing: Reconsidering complexity. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 79–103). Amsterdam: John Benjamins.

Cerezo, L., Manchón, R. M., & Nicolás-Conesa, F. (2019). What do learners notice while processing written corrective feedback? A look at depth of processing via written languaging. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 173–187). New York, NY: Routledge. https://doi.org/10.4324/9781315165080-12

Gass, S. (2010). Relating L2 input and L2 output. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 194–219). London: Continuum.

Hafner, C. A., Chik, A., & Jones, R. H. (2015). Digital literacies and language learning. *Language Learning & Technology*, 19(3), 1–7.

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11, 329–350. https://doi.org/10.1016/S1060-3743(02)00091-7

Hort, S. (2017). Exploring the use of mobile technologies and process logs in writing research. *International Journal of Qualitative Methods*. https://doi.org/10.1177/1609406917734060

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York, NY: Routledge. https://doi.org/10.4324/9781315887074

Leow, R. & Manchón, R. M. (Forthcoming). Expanding research agendas: Directions for future research agendas on writing, WCF, language learning, and ISLA. In R. M. Manchón & C. Polio (Eds.), *Handbook of SLA and writing*. New York, NY: Routledge.

López Serrano, S., Roca de Larios, J. & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically-motivated and empirically-based coding system. *Studies in Second Language Acquisition*, 41, 503–527. https://doi.org/10.1017/S0272263119000275

Manchón, R. M. (2014). The internal dimension of tasks: The interaction between task factors and learner factors in bringing about learning through writing. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning. Insights from and for L2 writing* (pp. 27–52). Amsterdam: John Benjamins.

Manchón, R. M. (2017). The potential impact of multimodal composition on language learning. *Journal of Second Language Writing*, 38, 94–95. https://doi.org/10.1016/j.jslw.2017.10.008

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge: CUP. https://doi.org/10.1017/9781108333603.015

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R. M. Manchón & P. K. Matsuda (Eds.), *The handbook of second and foreign language writing* (pp. 567–586). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-029

McKee, H. & DeVoss, D. N. (Eds.) (2007). *Digital writing research: Technologies, methodologies, and ethical issues*. New York, NY: Hampton Press.

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing*, 45, 31–45. https://doi.org/10.1016/j.jslw.2019.03.002

Murphy, L., & Roca de Larios, J. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL writers. *Journal of Second Language Writing*, 19(2), 61–8. https://doi.org/10.1016/j.jslw.2010.02.001

Oskoz, A., & Elola, I. (2020). *Digital L2 writing literacies. Directions for classroom practice*. London: Equinox.

Seror, J. (2013). Screen capture technology: A digital window into students' writing processes. *Canadian Journal of Learning and Technology*, 39(3), 1–16.

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Index

The current volume aspires to add to previous research on the connection between writing and language learning from a dual perspective: It seeks to reflect current progress in the domain as well as to foster future developments in theory and research. The theoretical postulations contained in Part I identify and expand in novel ways the diverse lenses through which the varied, multi-faceted dimensions of the connection between writing and language learning can be explored. The methodological reflections put forward in Part III signal theoretically-grounded and pedagogically-relevant paths along which future empirical work can grow. The empirical studies reported in Part II illuminate the myriad of individual, educational, and task-related variables that (may) mediate short-term and long-term language learning outcomes. These studies examine diverse forms of writing, performed in varied environments (including pen-and-paper and digital writing), conditions (writing individually and/or collaboratively), and instructional settings (academic settings – including secondary school and college level institutions – as well as out-of-school contexts).

John Benjamins Publishing Company