Bilingual Processing and Acquisition

12

Prediction in Second Language Processing and Learning

EDITED BY

Edith Kaan and Theres Grüter

John Benjamins Publishing Company

Prediction in Second Language Processing and Learning

Bilingual Processing and Acquisition (BPA)

ISSN 2352-0531

Psycholinguistic and neurocognitive approaches to bilingualism/multilingualism and language acquisition continue to gain momentum and uncover valuable findings explaining how multiple languages are represented in and processed by the human mind. With these intensified scholarly efforts come thought-provoking inquiries, pioneering findings, and new research directions. The *Bilingual Processing and Acquisition* book series seeks to provide a unified home, unlike any other, for this enterprise by providing a single forum and home for the highest-quality monographs and collective volumes related to language processing issues among multilinguals and learners of non-native languages. These volumes are authoritative works in their areas and should not only interest researchers and scholars investigating psycholinguistic and neurocognitive approaches to bilingualism/multilingualism and language acquisition but also appeal to professional practitioners and advanced undergraduate and graduate students.

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

Volume 12

Prediction in Second Language Processing and Learning
Edited by Edith Kaan and Theres Grüter

Prediction in Second Language Processing and Learning

Edited by

Edith Kaan
University of Florida

Theres Grüter
University of Hawaiʻi at Mānoa

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/bpa.12

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2021029295 (PRINT) / 2021029296 (E-BOOK)

ISBN 978 90 272 0970 2 (HB)
ISBN 978 90 272 5894 6 (E-BOOK)

John Benjamins Publishing Company · https://benjamins.com

Table of contents

Acknowledgments

This volume came together during the COVID-19 pandemic. We like to thank all contributors for their resilience and for their dedication to this volume in spite of so many other issues. We are also grateful to the following people for their constructive reviews: José Alemán Bañón, Arielle Borovsky, Carla Contemori, Monique Flecken, Adele Goldberg, Yi Ting Huang, Gerrit-Jan Kootstra, Sol Lago, Ryan Peters and Manuel Pulido.

List of contributors

Susanne Brouwer
Radboud University Nijmegen
Department of Modern Languages and Cultures
PO Box 9103
6500 HD Nijmegen
The Netherlands
s.brouwer@let.ru.nl

Angel Chan
Department of Chinese and Bilingual Studies
AG522, Core A
The Hong Kong Polytechnic University
Hung Hom
Hong Kong
angel.ws.chan@polyu.edu.hk

Si Chen
Department of Chinese and Bilingual Studies
AG522, Core A
The Hong Kong Polytechnic University
Hung Hom
Hong Kong
sarah.chen@polyu.edu.hk

Eunjin Chun
Department of Chinese and Bilingual Studies
AG522, Core A
The Hong Kong Polytechnic University
Hung Hom
Hong Kong
eunjin.chun@polyu.edu.hk

Claudia Felser
Potsdam Research Institute for Multilingualism
University of Potsdam
Karl-Liebknecht-Str. 24-25
14476 Potsdam
Germany
felser@uni-potsdam.de

Alice Foucart
Centro de Ciencia Cognitiva C3, Facultad de Lenguas y Educación
Universidad Nebrija
C. de Sta. Cruz de Marcenado, 27
28015 Madrid
Spain
afoucart@nebrija.es

Chiara Gambi
School of Psychology
Cardiff University
Tower Building
70 Park Place
CF10 3AT
Cardiff, Wales
U.K.
GambiC@cardiff.ac.uk

Theres Grüter
Department of Second Language Studies
University of Hawaiʻi at Mānoa
1890 East-West Road, Moore Hall, rm 570
Honolulu, HI 96822
USA
theres@hawaii.edu

Holger Hopp
Technische Universität Braunschweig
English Linguistics
Bienroder Weg 80
38106 Braunschweig
Germany
h.hopp@tu-braunschweig.de

Falk Huettig
Wundtlaan 1
6525 XD Nijmegen
The Netherlands
falk.huettig@mpi.nl

Aine Ito
Humboldt-Universität zu Berlin
Department of German Studies and Linguistics
Unter den Linden 6
10099 Berlin
Germany
aineito@gmail.com

Carrie N. Jackson
Dept of Germanic and Slavic Languages and Literatures
442 Burrowes
Penn State University
University Park, PA 16802
USA
cnj1@psu.edu

Edith Kaan
Department of Linguistics
University of Florida
4131 Turlington Hall
P.O. Box 115454
Gainesville, FL 32611
USA
kaan@ufl.edu

Figen Karaca
Radboud University Nijmegen
Erasmusplein 1
PO Box 9103
6500 HD Nijmegen
The Netherlands
f.karaca@let.ru.nl

Shulin Liu
Department of Chinese and Bilingual Studies
AG522, Core A
The Hong Kong Polytechnic University
Hung Hom
Hong Kong
shulin.liu@connect.polyu.hk

Martin J. Pickering
University of Edinburgh
Department of Psychology
7 George Square
Edinburgh EH8 9JZ
United Kingdom
martin.pickering@ed.ac.uk

Judith Schlenter
Department of Special Education and Rehabilitation
University of Cologne
Herbert-Lewin-Str. 10
50931 Cologne
Germany
judith.schlenter@gmail.com

Aleksandra Tomić
Institutt for språk og kultur
UiT Norges arktiske universitet
Postboks 6050 Langnes
9037 Tromsø
Norway
aleksandra.tomic@uit.no

Sharon Unsworth
Radboud University Nijmegen
Postbus 9103
6500 HD Nijmegen
The Netherlands
s.unsworth@let.ru.nl

Jorge R. Valdés Kroff
Department of Spanish and Portuguese Studies
University of Florida
PO Box 117405
170 Dauer Hall
Gainesville, Fl 32611
USA
jvaldeskroff@ufl.edu

Yanxin Alice Zhu
Department of Second Language Studies
University of Hawaiʻi at Mānoa
1890 East-West Road, Moore Hall, rm 570
Honolulu, HI 96822
USA
yanxinz@hawaii.edu

CHAPTER 1

Prediction in second language processing and learning

Advances and directions

Edith Kaan and Theres Grüter
University of Florida / University of Hawaiʻi at Mānoa

There is ample evidence that language users, including second language (L2) speakers, can predict upcoming information during listening and reading. Yet it is still unclear when, how, and why language users engage in prediction, and what the relation is between prediction and learning. This volume presents a collection of current research, insights and directions regarding the role of prediction in L2 processing and learning. In this introductory chapter, we provide an overview of the current state of the field and highlight that prediction may not always be the most efficient processing mechanism, depending on a language user's linguistic experience, task demands, goals and resources. We propose that a focus on the utility of prediction may help us better understand differences in predictive processing within and between individuals and groups.

Introduction

There is ample evidence that language users can predict upcoming information during listening and reading. For instance, many studies using visual world paradigms have shown that listeners can move their eyes to a particular object in the display even before the object is mentioned, suggesting they can develop clear expectations of how the sentence continues (Altmann & Kamide, 1999). According to some approaches, predictive processing is ubiquitous and essential for learning (Chang et al., 2006; Dell & Chang, 2014). Following the general trend in sentence processing research, predictive processing has become a popular topic in second language (L2) processing and acquisition research as well. While several initial L2 studies failed to find evidence for predictive processing in L2 learners (Grüter et al., 2012; Lew-Williams & Fernald, 2010; Martin et al., 2013), there now seems to be consensus that L2 users *can* predict. However, we are still in need of a better understanding of when, how, and why L2 (and first language!) comprehenders engage in prediction, and what the relation is between prediction and learning.

https://doi.org/10.1075/bpa.12.01kaa

The goal of this volume is to provide an overview of current research directions and recent insights related to prediction in L2 processing and learning. We will start this introductory chapter by addressing different views of prediction and its hypothesized role in language processing and learning. We then give a brief overview of past and current developments in predictive processing research in first and second language. We advocate an approach taking into account the utility of prediction (Kuperberg & Jaeger, 2016): Prediction may not always be the most efficient way to process language, depending on various factors. We believe that this approach may push research forward into explaining differences in predictive processing and learning within and across individuals and groups. We then summarize the chapters in this volume and conclude by sketching directions for further research.

What is prediction and what is it good for in L2 learning and processing?

What is prediction?

The term "prediction" has been interpreted in different ways in different research traditions in psycholinguistics and cognitive science. This makes the discussion about the function of predictive processing, and the evidence for prediction not always clear (DeLong et al., 2014; Kuperberg & Jaeger, 2016). A very wide conceptualization of "prediction" is that the "context influences the *state* of the language processing system *before* the bottom-up input is observed" (Kuperberg & Jaeger, 2016, p. 33). The "state of the language processing system" can be influenced by many factors, linguistic as well as non-linguistic, and changes with every new input. Aspects of the new input (semantics, form, and others) can therefore be more or less expected depending on the prior state of the processing system. As a result, the actual new input can be easier or harder to process.

In line with this wide view of prediction, Ferreira and Chantavarin (2018) propose the term "preparedness": a state of readiness for new information constrained by the context. According to Ferreira and Chantavarin, the function of predictive cues is to signal upcoming new information that is to some extent restricted by the preceding context (Myslín & Levy, 2016). This in turn will make it easier for the reader/listener to process (integrate) this new information. In this view, prediction and integration are two sides of the same coin. In addition, this stance implies that prediction of a very specific lexical form is rare in daily communication. The goal of communication is to provide new information. Words that are completely predictable do not convey new information, and hence will not occur often.

A more restrictive view of "prediction", commonly adopted in sentence processing, is that of pre-activation of various aspects of upcoming information. Given this concept of prediction, relevant research questions are: To what extent can specific words, including low-level features such as phonological or orthographic form, be pre-activated (e.g. Nieuwland, 2019)? Researchers adhering to this view of prediction hold that only effects observed *before* the onset of the critical word, or in very early stages of the processing of the target word, provide empirical evidence for prediction (Pickering & Gambi, 2018). Measures of processing difficulty at the target word or off-line measures are mostly inconclusive in this view, since such data can also be explained by re-active integration processes (Brouwer et al., 2017). Many researchers adhering to this view also distinguish pre-activation by (automatic, lexical) associations from more precise predictions constrained by the combination of various sources of top-down information (Ito & Pickering, this volume; Pickering & Gambi, 2018; Pickering & Garrod, 2013).

In sum, there is agreement that predictive processing involves some form of preparedness, expectancy or pre-activation within the language processing system. Yet views differ in whether this means that specific forms are pre-activated, and in what is taken to constitute evidence for prediction. What is uncontroversial is that language users can and do engage in proactive, forward-looking processing during language comprehension.

What is prediction good for?

Why do readers and listeners predict? Many assume that prediction makes communication more efficient because it makes upcoming information easier to process. The functions of these forward-looking aspects of communicative behavior have been of long-standing interest in the study of conversational turn-taking, following Sacks et al.'s (1974) observation that interlocutors predict, or project, the end of the current speaker's turn in preparation for their own response (see Levinson & Torreira, 2015, for discussion and psycholinguistic evidence).

Another function ascribed to predictive processing is its role in adaptation and learning. According to error-based implicit learning models (Chang et al., 2006), language users make predictions regarding upcoming information based on their prior experience. When these predictions are not borne out, the difference between the predicted and actual input ("prediction error") is then used to update the language processing system, such that future prediction errors can be minimized. Belief-updating models (Jaeger & Snider, 2013) are equivalent to error-based learning models in that the language processing system is changed (updated) with each incoming word in such a way that future changes are as small as possible. Error-based implicit learning and belief-updating can account for rapid adaptation by the listener and reader to, for instance, a particular way in which a particular person speaks, or to particular sentence structures used (Fine & Jaeger, 2013; Kleinschmidt et al., 2015). Error-based learning mechanisms have also been proposed to be a part of language learning (Chang et al., 2006; Phillips & Ehrenhofer, 2015). A learner makes predictions. If these predictions are not borne out, the difference between the predicted and actual input serves as implicit corrective feedback, pushing the learner more towards the target language. This proposal critically rest on the assumption that language learners can and do make predictions, and that they are able to use the implicit feedback provided by the actual input to update their processing system in case of a prediction error. These assumptions are not uncontested in the context of (L2) learning (Huettig & Mani, 2016; Kaan, 2015; Kaan et al., 2019; see also Hopp, this volume; Gambi, this volume).

Others have questioned the importance of prediction, pointing out that prediction is not always observed and is not essential for language processing and learning (Huettig & Mani, 2016). After all, most studies reporting evidence of prediction use highly constraining contexts, which may not be representative of natural language use. Moreover, there are phenomena that cannot be learned through prediction and there are mechanisms other than prediction through which learning occurs (Huettig & Mani, 2016; Kaan, 2015; see also Hopp, this volume). This suggests that prediction is not essential to processing and learning.

In this chapter we will suggest a way to reconcile the debate concerning the scope of prediction and its importance. We will start with a brief history of research on predictive processing.

A brief history of prediction in language processing research

Predictive processing has not always been a popular topic of research. In the 1980s and 1990s, modular conceptions of the mind typically stressed bottom-up flow of information (that is, from sensory perception to “higher” levels of cognition), with top-down information only serving as a way to select among multiple options provided by the bottom-up input. As a consequence, language processing research gave priority to bottom-up processing of perceptual and word-level information, with the role and timing of higher-level discourse and non-linguistic information being heavily debated. Predictive sentence processing was considered not very feasible: A sentence could end in an infinite number of possible ways, so predictions would often be wrong and therefore costly (see Ferreira & Chantavarin, 2018, for a concise historical review).

Around the first decade of the 21st century, cognitive science experienced a shift in thinking about information flow in the mind and brain. Friston (2010) and Clark (2013) argued for a view in which top-down information was of primary importance, with bottom-up information serving only a corrective role. In this view, the mind and brain are seen as "prediction machines" (Clark, 2013; Van Berkum, 2010). At any level of processing, predictions are made as to what will come next. These predictions are compared with what actually occurs. The difference between prediction and input is new information, which is used to fine-tune knowledge and processes at multiple levels. In this way, humans and other organisms can optimally adjust to the dynamic world around them.

Research on prediction in L1 sentence processing

In the realm of sentence processing, new experimental techniques such as visual world eyetracking (Tanenhaus et al., 1995), and event-related brain potentials (Kaan, 2007) provided evidence for predictive processing. Altmann and Kamide (1999) conducted an eye-tracking study showing that listeners can anticipate upcoming information: Listeners can move their eyes to a specific object in a display even before this object is verbally mentioned, that is, listeners look more at the edible object in the display when they hear *The boy will eat the…* compared to *The boy will move the….* Also results from studies using event-related brain potentials (ERPs) suggested that language users can predict upcoming words. A seminal study by DeLong et al. (2005) used highly constraining contexts such as *It was a windy day so the boy decided to fly a/an….* The ERPs showed incongruency effects at *a/an* when the form was not compatible with the most likely completion (e.g., when the form was *an* even though the most likely completion was *kite*). The size of the incongruency effect at the determiner correlated with how likely a completion the noun was. Note that these effects were seen before the noun was presented. This supported the view that prediction is gradual, and that language users predict upcoming words even to the level of the phonological form (that is, whether a word starts with a vowel or consonant, and hence requires *an* or *a*). Even though the replicability of these results is still being debated (Nicenboim et al., 2020; Nieuwland et al., 2018), studies using gender-marked languages such as Spanish and Dutch rather consistently found effects of incongruency when the gender marking on the determiner or adjective did not match that of an expected but yet-to-be-presented noun, suggesting that specific word features can be predicted, including gender information (Foucart et al., 2014; Foucart et al., 2016; Van Berkum et al., 2005; Wicha et al., 2004). Many studies conducted in recent years have shown that (adult L1) language users can predict, and can predict many types of information (lexical, thematic, phonological, syntactic, semantic, referential) on the basis on many types of cues (gender, case, discourse, prosody, verb form, verb meaning, among others); for a recent overview, see Pickering and Gambi (2018). Nevertheless, it remains under debate just how ubiquitous prediction is in adult L1 processing under various task demands and outside the laboratory (e.g. Huettig & Guerra, 2019; Pickering & Gambi, 2018).

Research on prediction in L2 sentence processing

Research on L2 sentence processing was initially mainly concerned with differences between L2 learners and native speakers and whether these differences could be accounted for by differences in representation (knowledge, e.g. Hawkins & Chan, 1997), processing and resources (Hopp, 2010), or both (e.g. Clahsen & Felser, 2006). Differences at the level of "processing" were conceived in terms of L2 users' potential limitations in incremental information *integration* during listening and reading (e.g. Sorace, 2011). However, following the growing interest in predictive processing in the early 2000s, studies soon appeared on predictive processing in L2 learners. Early studies (Grüter et al., 2012; Lew-Williams & Fernald, 2010; Martin et al., 2013) failed to find evidence for prediction in late second language learners. This led to views expressed in Kaan et al. (2010) and Grüter et al. (2014), hypothesizing that L2 learners have no or a Reduced Ability to Generate Expectations (RAGE). However, it quickly became clear that L2 speakers *can* make predictions during sentence comprehension. Even learners with very limited experience were shown to be able to use semantic information, such as verb restrictions, to anticipate a noun (Koehne & Crocker, 2014), and some L2 speakers were like native speakers in using gender information to predict upcoming nouns (Dussias et al., 2013; Hopp, 2013). It is therefore clear that the ability to predict in an L2 is not an all-or-nothing issue. In other words, the question is not whether L2 users can or cannot predict (they can), but what factors and circumstances affect the extent to which L2 users engage in predictive processing.

In a programmatic overview, Kaan (2014) outlined sources of variation in prediction. Many of these sources of variation are shared between native and L2 speakers (frequency, lexical quality, resources, task-induced processes and strategies, among others). Kaan (2014) therefore argued that L1 and L2 processing are not different in terms of predictive machinery, and that research should be aimed at identifying and quantifying sources of variation that drive predictive processing. Many studies on L2 prediction, including several in this volume, follow this line of inquiry.

A second, more recent line of research concerns the role of prediction in L2 learning. According to error-based learning models, prediction is a critical component of learning: If what you see or hear is different from what you expect, you will use this error to adjust your knowledge and future expectations. However, if L2 speakers engage in prediction less than or differently from native speakers, how exactly can learning from prediction errors contribute to L2 learning? This complex question has only recently begun to be addressed (Hopp, this volume; Jackson & Hopp, 2020; Kaan & Chun, 2018a).

In what follows, we will highlight some recent insights that we see as particularly relevant to predictive processing in L1 and L2 and its relation to learning, and advocate an approach that focuses on the utility of prediction (Kuperberg & Jaeger, 2016).

Prediction in L1 and L2 processing: The role of utility

Cue reliability and utility in L1 predictive processing

One major insight from L1 prediction research is that predictive processing is highly variable even *within* individuals, and that language users adapt their predictive behavior to particular situations, both as to *what* they are predicting and *whether* they are predicting. This ties in with observations and ideas concerning within-subject variability in other aspects of language processing (e.g. Ferreira & Patson, 2007; Hsu & Novick, 2016), and other cognitive functions (e.g. Botvinick et al., 2001; Kan et al., 2013). Prediction can be affected by many interrelated factors. We will discuss cue reliability, and goals and other task demands.

Predictions are made on the basis of cues. For instance, gender marking can be a cue for an upcoming type of noun; verb semantics can be cue for an upcoming type of object. Predictions are based on a reliable relation between the cue and what it signals in prior experience. Studies have shown that when cues are no longer reliable in the recent experimental context, participants stop using them predictively (Brothers et al., 2019; Heyselaar et al., 2020; Hopp, 2016; but see Zhang et al., 2019). For instance, if native German speaking participants are exposed to input that contains occasional errors in gender marking, they no longer use gender-marking as a predictive cue in the context of the study (Hopp, 2016). The reliability of a cue is also affected by the composition of the stimulus materials. For instance, when the proportion of related word pairs is high, participants show a larger and earlier relatedness effect in the ERPs than when the proportion is small, suggesting that participants engage more in prediction in the former case (Lau et al., 2013; see also Coulson et al., 1998; Hahne & Friederici, 2002 for comparable observations related to syntactic violations).

A cue can be considered more or less reliable in combination with other cues, or can change in what it is taken to be predictive of. Prosodic cues can enhance the predictive use of case markers in native German speakers (Henry et al., 2017). Listeners can change their predictions depending on who they are listening to and their prior experience with that talker (Kamide, 2012; Ryskin et al., 2020). Similarly, a talker's accent can affect predictive behavior. Listeners predict more slowly and to a lesser extent when listening to foreign accented than to native speech, especially when they have less experience listening to the foreign accent (Porretta et al., 2020; Romero-Rivas et al., 2016). This can be explained by the predictive cue being less reliable, either because it is harder or takes longer to perceive words in a foreign accent, or by the uncertainty of the predictions associated with the cue or speaker (see also Bosker et al., 2019).

The language user's goals and awareness can also affect prediction. Participants are more likely to predict when instructions explicitly direct them to do so (Brothers et al., 2017), and are more likely to make predictive eye-movements when they are aware of a predictive relation between the cue and the target (Andringa, 2020; Curcic et al., 2019). Another, related set of factors are task demands. In a visual world paradigm, participants are more likely to make predictive eye-movements if they have more time to look at the visual display in advance of the spoken information, and if the speech they listen to is presented at a slower rate (Ferreira et al., 2013; Huettig & Guerra, 2019). Ito et al. (2018) found that predictive eye-movements were delayed when participants had to memorize a set of unrelated words while listening to sentences. This suggests that predictive behavior can be modulated by concurrent memory load (see also Chun et al., this volume). Finally, predictive processing is affected by individual-difference factors such as reading ability, working memory, processing speed, age, and L1 exposure (Huettig & Brouwer, 2015; Huettig & Janse, 2015; Huettig & Pickering, 2019; Ito & Sakai, 2021; Mishra et al., 2012; Wlotko et al., 2012).

The above observations suggest that there is a lot of variability in prediction and in the way prediction occurs, even *within* individuals. Kuperberg and Jaeger (2016) therefore stress the role of utility: Language users weight the cost of prediction against the benefits to optimally reach their communicative goals. In some cases, using certain cues to predict will lead to too many errors or will exceed resources, and hence will be costly; in this case, depending on their goals, some listeners or readers may stop using this cue predictively, rely on a different cue, or may not predict at all in order to achieve optimal efficiency. Individuals therefore dynamically adjust their predictive behavior (what they predict and whether they predict) depending on the reliability of the predictive cues, their goals, task demands, resources, and other factors.

Cue reliability and utility in L2 predictive processing

As pointed out by Kaan (2014), the same sources that underlie variation in prediction in L1 may account for much of the variability we see in predictive processing in L2 learners and other bilinguals. Everything discussed in the previous section therefore applies to L2 learners as well. There are, however, some potential sources of variation that are unique to, or at least more pronounced, in L2 learners and bilinguals in general: cross-linguistic influence (competition or transfer between languages), proficiency, and exposure, including the way the L2 is learned. Under the view we advocate, these sources factor into cue reliability, cue weighting, and, consequently, the utility of prediction. The calculation of cost-benefit trade-offs may not always yield the same result in an L2 as in an L1. As a result, there may be circumstances where prediction is the most efficient processing option for L1 speakers but not for L2 speakers, or L2 speakers may make predictions that are different from those made by L1 speakers.

L2 users may place different weights on different cues, in line with the differential reliability of these cues within their language systems. According to the Competition Model (Bates & MacWhinney, 1981; MacWhinney & Bates, 1989) languages weigh cues differently. For instance, in English, word order is a reliable cue to determine the subject and object of a clause, whereas in Spanish subject-verb agreement is more important. What is a reliable cue for prediction in the L1 may more easily be regarded as a reliable cue in the L2 if these cues are shared between the languages. Evidence that reliable cues can be transferred from the L1 to the L2 comes from Van Bergen and Flecken (2017). In this study, L2 speakers of Dutch anticipated the orientation of an object based on the information encoded in the verb, but only if the learners' L1 also encoded object orientation in placement verbs. This suggests that semantic information encoded in the verb transfers from one language to the other and can be used predictively.

Morphosyntactic cues that overlap between L1 and L2 may also be used for prediction more, or more easily, in the L2. Dussias et al. (2013) report that intermediate-proficiency English L2 learners of Spanish did not use gender information predictively in Spanish, whereas proficiency-matched Italian learners of Spanish did, at least for feminine gender. This suggests that having a gender-marked L1 can help use gender predictively in a gender-marked L2 (Spanish). This especially holds when specific cues or constructions overlap (Hopp, this volume; Hopp & Lemmerth, 2018; Morales et al., 2015), suggesting that transfer of predictive use of a cue (e.g., gender) is not an all-not-nothing issue, but a function of what overlaps in the languages (see Foucart, this volume).

When there is no straightforward overlap in predictive use of cues between the languages, some cues may not be reliable for an L2 speaker because their representations are not sufficiently specified or entrenched. An example is the use of gender as a cue for a noun in beginning English L2 speakers of Spanish or German. If unreliable cues were used predictively, too many prediction errors would result. Hopp (2016) proposes that this situation is similar to that of L1 speakers who stop using gender predictively when they encounter a speaker who uses gender unsystematically. For an L2 learner, not predicting may be more efficient than generating predictions on the basis of unreliable cues.

Alternatively, L2 learners may interpret cues differently. For example, they may interpret a contrastive pitch accent as a lexical rather than an information-structural cue (Lee et al., 2020), or weigh cues differently, resulting in predictions that are different from those made by native speakers of the language. Grüter et al. (2020) investigated the predictive use of classifiers in Mandarin Chinese by L1 and L2 speakers. While L1 speakers relied on classifiers as a grammatical form class cue, L2 speakers also predicted upcoming nouns on the basis of classifiers but prioritized semantic cues, which are generally (but not perfectly) reliable. Another example is Hopp (2015), who tested English-speaking L2 learners of German with low-intermediate to advanced proficiency. Participants listened to sentences that started with an accusative or nominative noun phrase (e.g., *Den Wolf tötet gleich der Jäger.* 'the-ACC wolf kills soon the-NOM hunter', meaning 'The hunter will soon kill the wolf', and *Der Wolf tötet gleich den Hirsch*. 'the-NOM wolf kills soon the-ACC deer', meaning 'The wolf will soon kill the deer'). Native German speakers used the case information, reflecting the grammatical function and thematic status of a noun, at the start of the sentence to predict upcoming thematic roles: They predictively fixated on the most likely theme (deer in the example) when the first NP had nominative case, and on the most likely agent (hunter in the example) when the first NP was accusative. In contrast, the L2 speakers consistently made predictive eye-movements to the likely theme (deer), regardless of the case of the first noun phrase. These results can be accounted for in the following way: An agent-before-theme strategy is generally very reliable and yields the appropriate interpretation in most cases. Note that when tested off-line, L2 speakers typically perform correctly, so results cannot be explained by a lack of relevant linguistic knowledge. This suggests that in case of uncertainty (due to the processing speed needed when listening, uncertainty regarding what they hear, or due to less specified lexical representations), L2 speakers may rely more on cues they deem reliable from prior experience, that is, will mainly rely on frequent and prototypical associations, especially if these associations can be mapped from L1 to the L2.

Differential reliance on particular cues among L1 and L2 speakers may result from the different manner in which a language is learned in early childhood as opposed to later in life (Arnon & Ramscar, 2012; Grüter et al., 2012; Siegelman & Arnon, 2015). Take the case of grammatical gender. In the L2 classroom, nouns are typically first learned as form-meaning associations; rules pertaining to gender agreement, such as the selection of the correct determiner, are introduced later. By that time, the noun itself constitutes a fully reliable cue to meaning, thus blocking attention to co-occurrence relations with gender-marked determiners. On the other hand, a child acquiring language will hear nouns together with determiners and adjectives, and will form strong associations between the noun and gender information, hence making gender marking a reliable cue for the prediction of upcoming nouns. This account may in part also explain the difference in prediction on the basis of, e.g., gender between child L1 and L2 learners, and adult L2 learners (see also Karaca et al., this volume).

A note about proficiency

It is perhaps intuitive to assume that L2 speakers' engagement in predictive processing will become more similar to that of L1 speakers as global proficiency increases. Indeed, several studies have reported greater and/or more native-like effects of prediction in groups with higher global proficiency in the L2 (Chambers & Cooke, 2009; Dussias et al., 2013; Henry et al., 2020; Hopp, 2013; Hopp & Lemmerth, 2018; see also Karaca et al. this volume). Notably, however, a substantial number of recent studies that explicitly investigated the correlation between proficiency and prediction found no such effects (Dijkgraaf et al., 2017; Hopp, 2015; Ito et al., 2018; Kim & Grüter, 2020; Mitsugi, 2020; Perdomo & Kaan, 2019), or found that even very highly proficient L2 speakers differ from native speakers in their predictive use of certain features (Dijkgraaf et al., 2019; Kaan et al., 2016). Hopp (2015) report that proficiency did not affect the agent-verb-theme strategy (but cf. Henry et al., 2020). More proficient L2 speakers were however able to revise the initial interpretation more quickly than less proficient L2 speakers, indicating that the measure of proficiency employed in that study was able to explain relevant variation in processing behavior, but only with regard to information integration/revision, not prediction (see also Mitsugi, 2020). Although the absence of proficiency (or any other effect) needs to be interpreted with caution, observations such as the above suggest that prediction does not automatically come with global proficiency and, vice versa, that high overall proficiency does not imply L1-like prediction.

In sum, we propose that, as in L1 processing, prediction in L2 processing is modulated by utility: Generating specific predictions only makes sense in terms of processing efficiency if predictions are based on reliable cues and if predictions are correct most of the time. What is reliable is a function of experience and cross-linguistic influence, and the way the L2 is learned, but may not always be fully explained by global proficiency measures. Since L2 experience differs, it is a natural consequence that L2 speakers may differ from native speakers and from each other in what, when, and the extent to which they predict when processing the L2. Rather than constituting a deficiency, such differences may reflect adaptation to optimize processing efficiency given a language user's available knowledge, experience and resources.

Prediction and learning

Insights from L1 speakers

A second line of research featured in this volume concerns the role of prediction in (L2) learning. As we outlined above, error-based learning models assume that prediction plays a critical role in learning. The difference between what is predicted and the actual input is what gives rise to learning. That is, the processing system is changed as a result of prediction errors to minimize future prediction errors, and thus to more optimally process future input. Error-based learning mechanisms of the kind described have been successfully implemented in computational models of learning (Chang et al., 2006; Rumelhart et al., 1986). However, learning in humans is complicated, and listeners and readers do not exclusively or always use processing errors to adapt to the properties of the language they are exposed to.

One way in which error-based learning has been tested is by exposing participants to sentences with an unusual or infrequent syntactic structure. If readers or listeners adapt to these structures they will show a change in the way they process these structures as a function of exposure. Fine et al. (2013) report evidence indicating that L1 speakers adapt (see also, among others, Jaeger & Snider, 2013; Thothathiri & Snedeker, 2008; Wells et al., 2009). In Fine et al.'s study, participants read sentences with reduced relative clause ambiguities, such as *The experienced soldiers warned about the dangers conducted the midnight raid*. Fine et al. found the classic garden path effect, that is, reading times were longer at the disambiguating verb (*conducted*) for the temporarily ambiguous condition (example above) versus an unambiguous control sentence (*The experienced soldiers who were told about the dangers conducted the midnight raid*). Assuming a wide conceptualization of prediction, garden path effects can be interpreted as indexing prediction errors (Levy, 2008). Fine et al. found that the size of the garden path effect decreased over the course of the study, suggesting that readers adapted to the reduced relative clause structures. In addition, when two groups were compared, one with and one without prior exposure to reduced relatives, the group with prior exposure to reduced relatives showed smaller garden path effects. Finally, the group who read more reduced relatives eventually showed processing difficulty for simple main clause sentences, suggesting that these readers had reversed their processing preferences as a result of exposure. These findings support the view that participants adapt their processing on the basis of recent exposure. However, other adaptation studies yielded mixed results, with some studies or constructions showing evidence of adaptation but not others (see Kaan & Chun, 2018b for a critical review). For instance Kaan et al. (2019) found that native English speakers adapted to only one of two constructions tested. Hence even though (L1) readers experience prediction errors as reflected by the garden path effects, they do not always use these errors to adapt or adapt successfully to the structures they are exposed to in the timeframe of the study. The relation between prediction error and changes in the processing system is therefore not straightforward.

Insights from L2 speakers

Turning to L2 learning, learning approaches based on prediction error provide a seemingly elegant proposal for how L2 learning could occur through prediction. First, an L2 learner may base their prediction on what is reliable and frequent in their L1. If this prediction is not borne out in the L2, the L2 processing system will then adjust itself to minimize future prediction errors. This adjustment is larger the less expected the new input is. As a result, the learner's L2 processing system moves away from the L1 and starts to approximate the L2. Evidence supporting this comes from a priming study by Montero-Melis and Jaeger (2020), in which Swedish L2-learners of Spanish listened to directional descriptions and produced descriptions themselves. In Swedish, verbs mainly express the manner of the motion (as in e.g., *to push*), whereas Spanish verbs tend to encode the path of motion. Montero-Melis and Jaeger (2020) found that the less proficient L2 speakers adapted more quickly to the L2 structure that was infrequent in their L1 but frequent in the L2, whereas the more proficient speakers adapted more to the L2 structure that was infrequent in the L2. This can be accounted for in an error-based learning system in which beginner learners based themselves on the L1 regularities and adapted their system towards the L2, whereas the more advanced learners had already moved away from their L1 and fine-tuned the system to match the distributions of the structures in the L2.

However, the question is to what extent insights from such laboratory-based learning studies can inform our understanding of "learning" as typically understood in the context of L2 learning, namely the acquisition of long-term, generalizable knowledge (Kaan & Chun, 2018b). Jackson and colleagues (Jackson & Hopp, 2020; Jackson & Ruf, 2016) conducted a priming study and found that L2 learners adapted to structures that were infrequent in the L1 but frequent in the L2 (see also Kaan & Chun, 2018a). However, the effects disappeared in a post-test, suggesting that short-term adaptation did not lead to longer-term retention.

Another concern is that L2 speakers, like L1 speakers, do not always adapt with increased exposure, even though they may experience prediction errors. In a self-paced reading study, Kaan et al. (2019) found that both L2 and L1 speakers experienced garden-path effects when a sentence continued with an infrequent, unexpected structure. The L1 speakers showed a decrease of the garden path effect to one of the structures over the course of the study, suggesting they had changed their processing as a result of the recent exposure. However, the L2 speakers did not show signs of adaptation to this infrequent structure, suggesting that they did not use the prediction error to adapt, at least, not in the same timeframe (but cf. Hopp, 2020), even though this structure would have been even less frequent to the L2 speakers, and hence would be predicted to lead to more adaptation than in the L1 speakers. These results, on the face of it, are not what one would expect in an error-based learning approach.

Again: Utility

Why is it that L1 and L2 speakers sometimes do predict but do not adapt or learn? Hopp (this volume) discusses multiple reasons. Here we will focus on the utility of prediction. When a cue is very reliable for a speaker, there may be little or no gain to the speaker in making any changes to their existing system as a result of prediction error. In most cases, predictions will be correct and will contribute to communicative efficiency. Prediction errors induced by occasional exceptions will therefore have little or no effect on the reliability of a cue. In Bayesian terms: If prior information is very reliable, you will not change your beliefs much in the face of rare counterevidence (Kuperberg & Jaeger, 2016).

Aside from cue reliability and communicative efficiency, another factor that may enhance the utility of prediction, and as a consequence, adaptation, is attention and reward (Gambi, this volume; Hopp, this volume). An idea that has recently become more popular is that learners may need to experience "desirable difficulties" (Bjork & Kroll, 2015; Hertel, 2020). In short, learners need to be challenged and experience errors to learn. However, the challenges need to be such that they can potentially be overcome (hence "desirable"). Potts et al. (2019) show that learners who were asked to predict a translation (and hence, could make errors) before seeing the correct translation had better retention afterwards (see also Grüter et al., this volume). As Gambi (this volume) points out, learners who are asked to explicitly predict and sometimes make errors, may also pay more attention, and may be more motivated than learners in a more passive paradigm. For learners who are explicitly predicting, adaptation may therefore lead to a higher utility of the prediction, namely the reward of making a correct prediction.

Although controversial, there is some evidence that prediction and prediction errors can contribute to L2 learning, but that factors such as attention and awareness may play an important role. More research is needed to better understand what may induce L2 learners to predict and to learn from prediction errors, and how this could ultimately be implemented in the L2 classroom.

Concluding remarks, synopses of chapters, and future directions

Research on predictive language processing in the past decade has identified what kind of cues *can* be used to predict various kinds of upcoming information. The critical question is now how listeners and readers achieve maximal processing efficiency: Under what circumstances do listeners and readers use which cues to predict and adapt? And when is it more efficient not to predict? This approach shifts away from the debate over how essential or marginal prediction is for language processing and learning overall. Instead, it encourages research examining under what circumstances prediction is beneficial, and how these circumstances may vary between different language users and across development.

The chapters in this volume all contribute to this endeavor in various ways. **Ito** and **Pickering** offer a theoretical framing for the study of prediction in L2 processing within production-based models of prediction (Pickering & Garrod, 2013) and outline how the largely non-automatic processes that are hypothesized to underlie prediction-by-production may account for reduced engagement in prediction in an L2, especially in domains that align with later stages in the production process (syntactic and phonological encoding, vs. semantics).

Schlenter and **Felser** present an empirical contribution in which they examine L2 learners' ability to use information from different linguistic domains (lexical semantics, morphosyntax) for prediction during sentence processing in two eye-tracking studies with Russian learners of German. Their results show predictive use of both verbal semantics and case marking cues; yet subtle differences between L1 and L2 speakers emerged, with the latter showing greater uncertainty and susceptibility to competition, which manifested in delayed information integration.

Chun, Chen, Liu and **Chan** report novel empirical data from a visual-world eye-tracking study that extends the inquiry of semantic prediction to syntactically complex sentences containing relative clause attachment ambiguities. Their findings indicate that Chinese learners of English used semantic information from the verb to predict an upcoming object in both simple and complex syntactic contexts, yet prediction effects emerged somewhat later in the complex sentence condition, suggesting modulation of semantic prediction by the cognitive load associated with the concurrent processing of complex syntax.

Foucart's chapter offers a comprehensive review and discussion of the role that cross-linguistic influence (CLI) may play in L2 predictive processing, and outlines how CLI may be integrated into current models of prediction such as error-based implicit learning and prediction-by-production. **Karaca, Brouwer, Unsworth** and **Huettig** also address the role of CLI, yet focusing specifically on a population of language users that has received little attention in the prediction literature so far, bilingual children. Compared to monolingual children, bilingual children vary more extensively in the quantity and quality of the input they experience in each language, variability that is known to relate to the development of language skills in general, and predictive processing in particular, among monolingual children and adults. Bilingual children thus offer a unique opportunity for exploring the complex relationships between prediction, language proficiency, and language experience.

Tomić and **Valdés Kroff** address the role that code-switching may play as a predictive cue in bilingual processing. They summarize and discuss two studies in which Spanish-English bilinguals were presented with code-switched or unilingual sentences containing target words that differed in frequency or emotional valence. Findings indicate that experienced code-switchers can use a code switch event as a signal for upcoming information that is unexpected or unusual. These studies present a novel approach to the study of code-switching that shifts focus from the view of code-switching as a cause of integration cost to the benefits that a code-switch event might confer on processing downstream by way of prediction.

The last three chapters all focus on the relation between prediction and L2 learning. **Hopp** discusses the complex and reciprocal relation between predictive processing and learning and points to potential limitations in the role that prediction may play in the L2 learning of grammatical properties. As Hopp argues, learning from prediction error can only occur if the learner is able to revise a parse after the encounter of a prediction error. If either current knowledge or resources are insufficient to arrive at an appropriate revision, learning cannot take place, and unrevised prediction errors may result in persistent misparsing and fossilization. Hopp calls for future work assessing L2 learners' ability to benefit from prediction (error) across development and across tasks, and the role of awareness and explicit memory therein.

Drawing on insights from recent work on L1 vocabulary development in children and from research on novel word learning among adults, **Gambi** discusses the role that prediction may play in L2 vocabulary learning. Behavioral and neuroscientific evidence shows that guessing the meaning of an unknown word benefits the encoding of lexical knowledge in long-term memory. While these findings are consistent with error-driven learning accounts, Gambi argues that they are also compatible with explanations that point to the generation of a guess as a trigger for curiosity and motivation generated by the potential reward value of linguistic knowledge, and calls for future studies addressing the contributions of these factors in more naturalistic L2 word learning contexts.

In the last chapter, **Grüter, Zhu** and **Jackson** present findings from a structural priming experiment that examined whether forcing L2 learners to engage in prediction would lead to increased adaptation and learning. Their findings indicate that Korean L2 learners of English who had to guess how a virtual partner would describe pictures were more likely to use a construction used by that partner in their own subsequent picture descriptions than learners who only had to repeat the partner's description. Consistent with Gambi's discussion of guessing benefits in word learning, Grüter and colleagues observe that these results are consistent with error-driven learning accounts, but call for future work to examine the role of factors such as attention, awareness and motivation that may have contributed to the learning benefit observed when participants were forced to guess.

Collectively, the chapters in this volume present an impressive array of empirical and conceptual insights into the role of prediction in L2 processing and learning. Yet as they also make clear, we still have a long way to go to more fully understand the mutual relation between prediction and learning in L2. To push such research forward, we encourage longitudinal studies to investigate how prediction and its relation with learning develop over the course of learning, over the lifespan, and as a function of bilingual experience beyond global proficiency. Computational models are needed to better specify how multiple factors interact and operate, and when the cost of prediction outweighs its benefits (e.g. Aurnhammer & Frank, 2019; Kleinschmidt et al., 2015; Martin, 2016). Finally, research should be extended beyond traditional laboratory contexts, so as to test how current insights scale up to naturalistic conversations and real-life SLA contexts (Heyselaar et al., 2020; Schremm et al., 2017).

The goal of research on predictive processing is to account for differences within individuals, across individuals, as well as across populations. Specifying which cues are used to predict what, and determining when predicting is more efficient than not predicting, will be a big step towards that goal.

Acknowledgments

The authors would like to thank Holger Hopp for valuable feedback on an earlier draft.

Funding

EK is in part supported by NSF BSC-2017251.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Andringa, S. (2020). The emergence of awareness in uninstructed L2 learning: A visual world eye tracking study. *Second Language Research*, 36(3), 335–357. https://doi.org/10.1177/0267658320915502

Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, 122(3), 292–305. https://doi.org/10.1016/j.cognition.2011.10.009

Aurnhammer, C., & Frank, S. L. (2019). Evaluating information-theoretic measures of word prediction in naturalistic sentence reading. *Neuropsychologia*, 134, 107198. https://doi.org/10.1016/j.neuropsychologia.2019.107198

Bates, E., & MacWhinney, B. (1981). Second language acquisition from a functionalist perspective: Pragmatic, semantic and perceptual strategies. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences conference on native and foreign language acquisition* (pp. 190–214). New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.1981.tb42009.x

Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *The American Journal of Psychology*, 128(2), 241–252. https://doi.org/10.5406/amerjpsyc.128.2.0241

Bosker, H. R., van Os, M., Does, R., & van Bergen, G. (2019). Counting 'uhm's: How tracking the distribution of native and non-native disfluencies influences online language comprehension. *Journal of Memory and Language*, 106, 189–202. https://doi.org/10.1016/j.jml.2019.02.006

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Brothers, T., Dave, S., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2019). Flexible predictions during listening comprehension: Speaker reliability affects anticipatory processes. *Neuropsychologia*, 135, 107225. https://doi.org/10.1016/j.neuropsychologia.2019.107225

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, 93, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

Brouwer, S., Sprenger, S., & Unsworth, S. (2017). Processing grammatical gender in Dutch: Evidence from eye movements. *Journal of Experimental Child Psychology*, 159, 50–65. https://doi.org/10.1016/j.jecp.2017.01.007

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35(4), 1029–1040. https://doi.org/10.1037/a0015901

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42. https://doi.org/10.1017/S0142716406060024

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language & Cognitive Processes*, 13(1), 21–58. https://doi.org/10.1080/016909698386582

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, 69(S1), 42–71. https://doi.org/10.1111/lang.12321

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1634). https://doi.org/10.1098/rstb.2012.0394

DeLong, K. A., Troyer, M., & Kutas, M. (2014). Pre-processing in sentence comprehension: Sensitivity to likely upcoming meaning and structure. *Language and Linguistics Compass*, 8(12), 631–645. https://doi.org/10.1111/lnc3.12093

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35, 353–387. https://doi.org/10.1017/S0272263112000915

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, 27(6), 443–448. https://doi.org/10.1177/0963721418794491

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, 69(3), 165–182. https://doi.org/10.1016/j.jml.2013.06.001

Ferreira, F., & Patson, N. D. (2007). The 'Good Enough' approach to language comprehension. *Language and Linguistics Compass*, 1(1–2), 71–83. https://doi.org/10.1111/j.1749-818X.2007.00007.x

Fine, A. B., & Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, 37(3), 578–591. https://doi.org/10.1111/cogs.12022

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, 8(10), 1–18. https://doi.org/10.1371/journal.pone.0077661

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 4161–1469. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2016). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 19(1), 213–219. https://doi.org/10.1017/S1366728915000486

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. https://doi.org/10.1038/nrn2787

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, 35(2), 221–234. https://doi.org/10.1080/23273798.2019.1648840

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191–215. https://doi.org/10.1177/0267658312437990

Grüter, T., Rohde, H., & Shafer, A. J. (2014). The role of discourse-level expectations in non-native speakers' referential choices. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the 38th Annual Boston University Conference on Language Development* (pp. 179–191). Cascadilla Press.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, 13, 339–356. https://doi.org/10.1016/S0926-6410(01)00127-6

Hawkins, R., & Chan, C. Y.-H. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, 13(3), 187–226. https://doi.org/10.1191/026765897671476153

Henry, N., Hopp, H., & Jackson, C. N. (2017). Cue additivity and adaptivity in predictive processing. *Language, Cognition and Neuroscience*, 32(10), 1229–1249. https://doi.org/10.1080/23273798.2017.1327080

Henry, N., Jackson, C. N., & Hopp, H. (2020). Cue coalitions and additivity in predictive processing: The interaction between case and prosody in L2 German. *Second Language Research*, 0267658320963151. https://doi.org/10.1177/0267658320963151

Hertel, P. (2020). Everyday challenges to the practice of desirable difficulties: introduction to the forum. *Journal of Applied Research in Memory and Cognition*, 9(4), 425–427. https://doi.org/10.1016/j.jarmac.2020.09.002

Heyselaar, E., Peeters, D., & Hagoort, P. (2020). Do we predict upcoming speech content in naturalistic environments? *Language, Cognition and Neuroscience*, 1–22. https://doi.org/10.1080/23273798.2020.1859568

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120, 901–931. https://doi.org/10.1016/j.lingua.2009.06.004

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32(2), 277–307. https://doi.org/10.1177/0267658315624960

Hopp, H. (2020). Morphosyntactic adaptation in adult L2 processing: Exposure and the processing of case and tense violations. *Applied Psycholinguistics*, 41(3), 627–656. https://doi.org/10.1017/S0142716420000119

Hopp, H., & Lemmerth, N. (2018). Lexical and syntactic congruency in l2 predictive gender processing. *Studies in Second Language Acquisition*, 40(1), 171–199. https://doi.org/10.1017/S0272263116000437

Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science*, 27(4), 572–582. https://doi.org/10.1177/0956797615625223

Huettig, F., & Brouwer, S. (2015). Delayed anticipatory spoken language processing in adults with dyslexia – Evidence from eye-tracking. *Dyslexia*, 21(2), 97–122. https://doi.org/10.1002/dys.1497

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research*, 1706, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013

Huettig, F., & Janse, E. (2015). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, 31(1), 19–31. https://doi.org/10.1080/23273798.2015.1072223

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences*, 23(6), 464–475. https://doi.org/10.1016/j.tics.2019.03.008

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., & Sakai, H. (2021). Everyday language exposure shapes prediction of specific words in listening comprehension: A visual world eye-tracking study [Original Research]. *Frontiers in Psychology*, 12(240). https://doi.org/10.3389/fpsyg.2021.607474

Jackson, C. N., & Hopp, H. (2020). Prediction error and implicit learning in L1 and L2 syntactic priming. *International Journal of Bilingualism*, 24(5–6), 895–911. https://doi.org/10.1177/1367006920902851

Jackson, C. N., & Ruf, H. T. (2016). The priming of word order in second language German. *Applied Psycholinguistics*, 38(2), 315–345. https://doi.org/10.1017/S0142716416000205

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and Linguistics Compass*, 1(6), 571–591. https://doi.org/10.1111/j.1749-818X.2007.00037.x

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E. (2015). Knowing without predicting, predicting without learning. *Linguistic Approaches to Bilingualism*, 5(4), 482–486. https://doi.org/10.1075/lab.5.4.07kaa

Kaan, E., & Chun, E. (2018a). Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, 21, 228–242. https://doi.org/10.1017/S1366728916001231

Kaan, E., & Chun, E. (2018b). Syntactic adaptation. In K. D. Federmeier & D. G. Watson (Eds.), *Psychology of learning and motivation* (Vol. 68, pp. 85–116). Academic Press. https://doi.org/10.1016/bs.plm.2018.08.003

Kaan, E., Dallas, A. C., & Wijnen, F. (2010). Syntactic predictions in second-language sentence processing. In J.-W. Zwart & M. de Vries (Eds.), *Structure preserved. Festschrift in the honor of Jan Koster* (pp. 207–213). John Benjamins. https://doi.org/10.1075/la.164.23kaa

Kaan, E., Futch, C., Fernández Fuertes, R., Mujcinovic, S., & Álvarez de la Fuente, E. (2019). Adaptation to syntactic structures in native and non-native sentence comprehension. *Applied Psycholinguistics*, 40(1), 3–27. https://doi.org/10.1017/S0142716418000437

Kaan, E., Kirkham, J., & Wijnen, F. (2016). Prediction and integration in native and second-language processing of elliptical structures. *Bilingualism: Language and Cognition*, 19(1), 1–18. https://doi.org/10.1017/S1366728914000844

Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition*, 124(1), 66–71. https://doi.org/10.1016/j.cognition.2012.03.001

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, 129(3), 637–651. https://doi.org/10.1016/j.cognition.2013.09.001

Kim, H., & Grüter, T. (2020). Predictive processing of implicit causality in a second language: A visual-world eye-tracking study. *Studies in Second Language Acquisition*, 1–22. https://doi.org/10.1017/S0272263120000443

Kleinschmidt, D. F., Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, 122(2), 148–203. https://doi.org/10.1037/a0038695

Koehne, J., & Crocker, M. W. (2014). The interplay of cross-situational word learning and sentence-level constraints. *Cognitive Science*, 39(5), 849–889. https://doi.org/10.1111/cogs.12178

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition & Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lau, E., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single word contexts. *Journal of Cognitive Neuroscience*, 25(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Lee, A., Perdomo, M., Kaan, E. (2020). Native and second-language processing of contrastive pitch accent: An ERP study. *Second Language Research*, 36(4), 503–527. https://doi.org/10.1177/0267658319838300

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language [Original Research]. *Frontiers in Psychology*, 6(731). https://doi.org/10.3389/fpsyg.2015.00731

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003

MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing*. Cambridge University Press.

Martin, A. E. (2016). Language processing as cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in Psychology*, 7(120). https://doi.org/10.3389/fpsyg.2016.00120

Martin, C., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eyemovements are modulated by reading ability – Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1), 1–10. https://doi.org/10.16910/jemr.5.1.3

Mitsugi, S. (2020). Generating predictions based on semantic categories in a second language: A case of numeral classifiers in Japanese. *International Review of Applied Linguistics in Language Teaching*, 58(3), 323–349. https://doi.org/10.1515/iral-2017-0118

Montero-Melis, G., & Jaeger, T. F. (2020). Changing expectations mediate adaptation in L2 production. *Bilingualism: Language and Cognition*, 23(3), 602–617. https://doi.org/10.1017/S1366728919000506

Morales, L., Paolieri, D., Dussias, P. E., Valdés Kroff, J. R., Gerfen, C., & Bajo, M. T. (2015). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition*, 19(2), 294–310. https://doi.org/10.1017/S1366728915000176
Myslín, M., & Levy, R. (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, 147, 29–56. https://doi.org/10.1016/j.cognition.2015.10.021
Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a Bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, 142, 107427. https://doi.org/10.1016/j.neuropsychologia.2020.107427
Nieuwland, M. S. (2019). Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews*, 96, 367–400. https://doi.org/10.1016/j.neubiorev.2018.11.019
Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., Donaldson, D. I., Kohút, Z., Rueschemeyer, S. A., & Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *Elife*, 7. https://doi.org/10.7554/eLife.33468
Perdomo, M., & Kaan, E. (2019). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research*. https://doi.org/10.1177/0267658319879196
Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, 5(4), 409–453. https://doi.org/10.1075/lab.5.4.01phi
Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158
Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. https://doi.org/10.1017/S0140525X12001495
Porretta, V., Buchanan, L., & Järvikivi, J. (2020). When processing costs impact predictive processing: The case of foreign-accented speech and accent experience. *Attention, Perception, & Psychophysics*, 82(4), 1558–1565. https://doi.org/10.3758/s13414-019-01946-7
Potts, R., Davies, G., & Shanks, D. R. (2019). The benefit of generating errors during learning: What is the locus of the effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(6), 1023–1041. https://doi.org/10.1037/xlm0000637
Romero-Rivas, C., Martin, C. D., & Costa, A. (2016). Foreign-accented speech modulates linguistic anticipatory processes. *Neuropsychologia*, 85, 245–255. https://doi.org/10.1016/j.neuropsychologia.2016.03.022
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. https://doi.org/10.1038/323533a0
Ryskin, R., Ng, S., Mimnaugh, K., Brown-Schmidt, S., & Federmeier, K. D. (2020). Talker-specific predictions during language processing. *Language, Cognition and Neuroscience*, 35(6), 797–812. https://doi.org/10.1080/23273798.2019.1630654
Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735. https://doi.org/10.1353/lan.1974.0010

Schremm, A., Hed, A., Horne, M., & Roll, M. (2017). Training predictive L2 processing with a digital game: Prototype promotes acquisition of anticipatory use of tone-suffix associations. *Computers & Education*, 114, 206–221. https://doi.org/10.1016/j.compedu.2017.07.006

Siegelman, N., & Arnon, I. (2015). The advantage of starting big: Learning from unsegmented input facilitates mastery of grammatical gender in an artificial language. *Journal of Memory and Language*, 85, 60–75. https://doi.org/10.1016/j.jml.2015.07.003

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33. https://doi.org/10.1075/lab.1.1.01sor

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. *Cognition*, 108(1), 51–68. https://doi.org/10.1016/j.cognition.2007.12.012

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad – Any implications for neuropragmatics? *Italian Journal of Linguistics*, 22, 181–208.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–467. https://doi.org/10.1037/0278-7393.31.3.443

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, 58(2), 250–271. https://doi.org/10.1016/j.cogpsych.2008.08.002

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in spanish sentence reading. *Journal of Cognitive Neuroscience*, 16(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*. https://doi.org/10.1037/a0029206

Zhang, W., Chow, W.-Y., Liang, B., & Wang, S. (2019). Robust effects of predictability across experimental contexts: Evidence from event-related potentials. *Neuropsychologia*, 134, 107229. https://doi.org/10.1016/j.neuropsychologia.2019.107229

CHAPTER 2

Automaticity and prediction in non-native language comprehension

Aine Ito and Martin J. Pickering
Humboldt-Universität zu Berlin / University of Edinburgh

Some evidence suggests that prediction is more limited in non-native language (L2) than native language (L1) comprehension. We evaluate the hypothesis that prediction is limited in L2 because prediction is largely non-automatic. We examine whether the subprocesses involved in prediction are *unconscious, unintentional, efficient* and *uncontrollable* (Bargh, 1994) to understand the extent to which prediction is automatic in L1 and L2. To unpack the subprocesses in prediction, we draw on Pickering and Garrod's (2013) proposal that people primarily use their production system for prediction, as well as a more automatic association-based mechanism. We conclude that at least some of the subprocesses in prediction are not fully automatic and suggest that these non-automatic processes can interfere with prediction in L2.

Introduction

People regularly predict upcoming information during language comprehension (Pickering & Gambi, 2018). By prediction in language comprehension, we mean pre-activation of information (e.g., semantic, syntactic, and phonological components of lexical representations) of an input (e.g., word, phrase) before the comprehender hears or reads the input. For example, upon hearing the verb *eat*, people may pre-activate the semantic representation [+edible] about an upcoming object. When a specific word (e.g., *cake*) is highly predictable, they may pre-activate the phonological (e.g., /keɪk/) or syntactic (e.g., a noun) components of that word's lexical representation. Such predictions likely facilitate efficient sentence processing by allowing people to access and process the information that has not yet occurred in the sentence. However, L2 speakers often predict more slowly or predict less detailed information than L1 speakers, as we review below. This lesser degree of prediction during comprehension might partly explain why L2 speakers have difficulty with online sentence processing. Thus, it is important to understand the

https://doi.org/10.1075/bpa.12.02ito

mechanisms of prediction in L2 comprehension and how they might differ from those in L1 comprehension.

One possible explanation for the limitation in L2 prediction is that it is largely non-automatic. More generally, prediction is thought to involve both automatic processes (which are rapid and not subject to resource limitation) and non-automatic processes (which require time and resources) (Huettig, 2015; Pickering & Garrod, 2013). Pickering and Garrod (2013; see also Pickering & Gambi, 2018) proposed that the main mechanism people use for prediction is the mechanism they use to produce language. Since language production involves non-automaticity at each critical stage (conceptualization, lexical selection and grammatical encoding), prediction using the production system is also likely to be largely non-automatic. If so, L2 speakers may not have enough cognitive resources available for prediction, because L2 processing is generally more resource-demanding than L1 processing (Segalowitz & Hulstijn, 2009). We evaluate this hypothesis by examining the extent to which prediction during L2 processing is automatic.

This chapter falls into four sections. In the first section, we consider a graded view of automaticity (i.e., rather than all or none) in relation to "the four horsemen of automaticity" proposed by Bargh (1994). In the second section, we apply this notion to a model of language prediction. In the third and fourth sections, we review studies on prediction in L1 and L2 respectively, and reconcile their findings with the production-based prediction model (Pickering & Gambi, 2018; Pickering & Garrod, 2013).

Graded view of automaticity

Traditionally, automaticity has been seen as dichotomous: a process is either fully automatic or fully non-automatic (Shiffrin & Schneider, 1977). Automatic processes were considered to be unconscious, unintentional, efficient, and uncontrollable. Non-automatic processes were simply the opposite: conscious, intentional, inefficient, and controllable. However, this notion has been challenged. Some researchers (Bargh, 1994; Moors, 2016) have argued that processes are more or less automatic rather than fully automatic or non-automatic, because the features of automaticity are not necessarily co-present. That is, a process can be automatic in some respects but not in others. For example, a process can be uncontrollable but conscious. Thus, automaticity is seen as a graded notion according to this more recent view.

Bargh (1994) decomposed automaticity into four features, which he called "four horsemen": *awareness, intentionality, efficiency*, and *controllability* (see Garrod & Pickering, 2007, for an application of this account to language production). *Awareness* refers to whether people are aware of the stimulus itself or its influence.

Intentionality refers to whether people can instigate the process intentionally. *Efficiency* refers to whether the process draws on attentional resources. *Controllability* refers to whether people can stop or moderate the process once it has started. A process is regarded as most automatic if it occurs unconsciously, demands no attentional resources, and cannot be intentionally started or stopped. For example, masked priming is largely automatic in these respects, as participants are unaware of the stimuli (primes) and the priming effect occurs regardless of participants' intention. As we review below, much of the evidence suggesting that prediction is non-automatic relates to efficiency, with some evidence suggesting that it is also non-automatic in terms of awareness, intentionality, and controllability.

A production-based model of prediction

In this section, we apply the four horsemen of automaticity proposed by Bargh (1994) to the mechanisms of prediction in the production-based prediction model (Pickering & Gambi, 2018; Pickering & Garrod, 2007, 2013). This model proposes that the central mechanism of prediction is *prediction-by-production*: people predict upcoming language with the mechanism they use to produce language. To do so, they covertly imitate the language they have comprehended, and derive the intention of the speaker while considering both linguistic and non-linguistic information (e.g., shared background knowledge). They may additionally apply adjustments to compensate for differences between the speaker and themselves, which allows them to predict what the speaker, rather than themselves, is likely to say. They then run the speaker's intention through their own production system, retrieving the representations of what constitutes the upcoming utterance of the speaker. Prediction-by-production is relatively slow and requires resources because critical stages of language production require time and resources. People can also control and are aware of at least some of the production processes (e.g., lexical selection). Thus, prediction-by-production appears to involve processes that are non-automatic in terms of efficiency, awareness, intentionality, and controllability.

There is a complication to this account. Pickering and Garrod (2013) proposed that there is another aspect of prediction-by-production that is more automatic. This aspect of prediction involves a so-called forward model (see Wolpert, 1997). When speakers plan an utterance (e.g., *I will eat a cake*.), they also rapidly predict aspects of what they are going to say (e.g., the first phoneme /k/) using a forward model. This model involves a (learned) mapping between the speaker's intention and the predicted outcome. In the past, whenever the speaker has decided to say *cake*, she has subsequently uttered /*k*/ and experienced the sensation of uttering it, so the decision to utter *cake* now activates this sensation automatically. This prediction is used in self-monitoring: if the speaker now prepares the wrong phoneme (e.g., /*j*/), she notices that it mismatches the forward model and can revise her utterance.

When the comprehender predicts an utterance using prediction-by-production, he can also use the forward model, but this time using his covert imitation of the speaker's utterance (plus adjustments for differences between comprehender and speaker). After hearing *I will eat a*, he constructs the intention to produce the word *cake*, and at the same time predicts (say) /*k*/ via a forward model. Prediction using the forward model is highly automatized.

But at the same time, he begins to implement the process of producing the word *cake* and thereby generates predictions of what the speaker will say (Pickering & Gambi, 2018) – predicting its meaning, grammar, and sound in turn. This approach to prediction is slower and resource-intensive (as it is implemented in the language production system). For the purposes of this chapter, we focus on this latter aspect of prediction-by-production – in which the comprehender computes representations corresponding to the stages of language production (see Pickering & Gambi, 2018).

Prediction-by-production is complemented by another mechanism called *prediction-by-association*, which is not based on production. In this mechanism, lexical priming plays a role, and representations that are associated with words present in the context receive activation. It appears to be automatic in terms of efficiency and awareness, because lexical priming occurs very rapidly (Ratcliff & McKoon, 1981), and people do not seem to be aware of the process. It appears largely automatic in terms of controllability and intentionality as well (e.g., because people cannot stop it). There is some evidence that associations-based priming effects might be affected by strategic processing (Lau et al., 2013) but not at early stages (Grossi, 2006), suggesting that rapid spreading of activation is uncontrollable. While prediction-by-association is more rapid and consumes fewer resources, it is not as accurate as prediction-by-production, because words in the context often have lexical associations with words that are unlikely to follow the context. For example, the context *I will ride…* is likely to activate words representing a vehicle, but words like *motorbike* are not likely to occur if the speaker is a child. People take such information into account when predicting upcoming language (Kamide et al., 2003), but prediction-by-association activates lexically associated words even if they do not fit with the context. Thus, Kukona et al. (2011) found that when people heard *Toby will arrest…*, they tended to look at a correct theme (e.g., crook) as well as a typical agent of the verb (e.g., policeman) even though it was not a plausible object of the verb (see the next section for details of this study).

Automaticity of prediction in L1

We propose that prediction-by-production is largely non-automatic, as critical stages in language production involve non-automatic processes (Hartsuiker & Moors, 2017; Roelofs & Piai, 2011) – though note the discussion of forward modelling above. Language production involves several steps from conceptualization of the meaning that a speaker wants to convey, selection of a lexical item, morphological encoding, and retrieval of the phonological form up to articulation. Conceptualization is cognitively demanding (Bock, 1982), people are aware of the process, and they can decide to start or interrupt the process; hence it is largely non-automatic in terms of all four horsemen (Garrod & Pickering, 2007). Lexical selection seems non-automatic in terms of awareness and intentionality, as speakers are generally aware of their lexical choices and they can voluntarily instigate the process (Garrod & Pickering, 2007).

Processes of linguistic encoding such as syntactic planning, phonological word-form selection, and phoneme selection also seem to require working memory (e.g., Cook & Meyer, 2008; Ferreira & Pashler, 2002; Hartsuiker & Barkhuysen, 2006), and hence are non-automatic in terms of the horseman of efficiency. For example, Hartsuiker and Barkhuysen (2006) found more number agreement errors (e.g., *The coupon in the flyers were…*) when participants had to remember a word list than when they did not, suggesting that syntactic planning in speech production requires working memory resources and is therefore not fully automatic (see also Fayol et al., 1994). Experiments using a dual-task (picture naming and tone discrimination) found that picture naming latency was shorter when it followed a high versus low cloze context, when the lexical frequency was high versus low (Ferreira & Pashler, 2002), and when a distractor presented with the target picture was phonologically related to the target than when it was phonologically unrelated (Cook & Meyer, 2008). Critically, these effects carried forward to the following tone discrimination task, suggesting that lemma selection (affected by cloze), word-form selection (affected by frequency), and phoneme selection (affected by phonological relatedness) all require cognitive resources (because, if not, they should not affect the tone-discrimination task) and are non-automatic in terms of efficiency. Some subprocesses are treated as automatic in some production models, such as spreading activation between lemma and lexeme (Dell, 1986) or formulating and articulatory procedures (Levelt, 1989). But critically, the above findings suggest that each key stage of production involves some non-automaticity.

Similarly, there is evidence that prediction is not fully automatic in terms of efficiency, and whether prediction occurs is subject to time and resources available during comprehension. For example, Huettig and Janse (2016) investigated the effect of people's working memory span and general processing speed on their predictive eye movements using the visual world paradigm, where people listen to language while viewing some objects. In their study, L1 Dutch speakers listened to Dutch sentences such as *Kijk naar de afgebeelde piano* ('Look at the$_{\text{-common}}$ displayed piano$_{\text{-common}}$.') while viewing a scene containing four objects. Articles in Dutch are gender-marked (e.g., *de* is common gender), and the gender of only the target object name matched the article's gender. As expected, participants fixated the target object predictively (i.e., before the target noun was mentioned). Critically, their working memory and processing speed (measured using multiple verbal and non-verbal tasks) were positively related to the degree of predictive eye movements: Participants who had a greater working memory span and a higher processing speed made more predictive eye movements. These effects suggest that some processes involved in making predictive eye movements require resources, and hence are non-automatic with respect to efficiency in the four horsemen.

The findings in Huettig and Janse (2016) are strengthened by evidence showing a causal link between working memory and predictive eye movements. Ito, Corley et al. (2018; Experiment 1) found that an additional working memory load interfered with predictive eye movements. L1 English speakers listened to predictable sentences (i.e., with a constraining verb) such as *The lady will fold the scarf* or unpredictable sentences (i.e., with a less constraining verb) such as *The lady will find the scarf* while viewing a scene with four objects. Only the target object (here, scarf) was plausible after the verb in predictable sentences (and hence it was predictable given that one of the pictures would be described), whereas all four objects were plausible after the verb in unpredictable sentences. The participants' task was to click on the mentioned object. Half of the participants performed an additional working memory task to tax their working memory (i.e., to make fewer cognitive resources available) during sentence processing. They saw five abstract words before the listen-and-click task and recalled the words immediately after the task. Participants fixated the target object predictively in predictable sentences, but these predictive eye movements were delayed in those who performed the working memory task compared to those who did not. This finding, together with the findings in Huettig and Janse (2016), suggests that cognitive resources are required for making predictive eye movements.

There is also evidence that prediction is dependent on time available for generating predictions, which also suggests that prediction is non-automatic in terms of efficiency. Huettig and Guerra (2019) used the scenes and sentences from Huettig and Janse (2016) and investigated effects of speech rate and preview time on predictive eye movements. When participants viewed the objects for four seconds before the sentence began, they fixated the target predictively both when the speech was at a normal rate (the sentence took just under 2 seconds) and at a slow rate (it took just over 4 seconds). However, when they were given a one-second preview, predictive eye movements occurred only in the slow rate condition. With the shorter preview, it could be that participants were still in the process of encoding of the scene, which then interfered with generating predictions.

Similarly, Ito et al. (2016) found that people do not pre-activate full representations of a predictable word when they do not have enough time to generate predictions. L1 English speakers read predictive sentence contexts (e.g., *The student is going to the library to borrow a…*), followed by a predictable word (*book*), a word that was form-related (*hook*) or semantically related (*page*) to the predictable word, or an unrelated word (*sofa*). Participants read the sentences word-by-word at a 500 ms or 700 ms presentation rate. Consistent with previous studies (Federmeier & Kutas, 1999; Laszlo & Federmeier, 2009), Ito et al. found that all unexpected words elicited an N400 effect relative to the predictable word, but words that were form-related or semantically related to the predictable word elicited a reduced N400 effect relative to the unrelated word. The N400 reduction was greater when the word predictability was higher (estimated in a cloze test, where people completed sentence fragments with a continuation that came to mind), suggesting that people pre-activated word form and meaning of the predictable word. Critically, Ito et al. found the N400 reduction for semantically related words at both presentation rates, but the N400 reduction for form-related words emerged only at the slower presentation rate (700 ms). This finding suggests that prediction of the phonological component of lexical representations takes time; when time is limited, the semantic component may be predicted but not the phonological component.

The effects of working memory and time on prediction suggest that some processes involved in prediction are not automatic in terms of efficiency. But which predictive processes require time and cognitive resources? There is evidence indicating that adjusting the perspective to the speaker in prediction may be cognitively demanding. Corps et al. (2021) tested prediction of an upcoming object in sentences such as *I would like to wear the nice … tie/dress*, where the continuation is likely to be different depending on the gender of the speaker (e.g., a male speaker might say *tie*, whereas a female speaker might say *dress*). That is, the listener must adjust for differences in perspective to predict what the speaker is likely to say. L1 English speakers heard these sentences in a male or female voice (the continuation always matched the speaker's gender) while viewing objects that did or did not match the speaker's gender (e.g., a tie, a dress) together with two distractors (e.g., a drill, a hairdryer). Upon hearing the verb, they fixated objects that were compatible with the verb (e.g., a tie, a dress) over the distractors irrespective of the speaker's gender. They also predictively fixated objects that were expected based on the speaker's gender (e.g., a tie for a male speaker), but these eye movements occurred later than association-based predictive eye movements. The delayed effect for speaker-specific prediction suggests that adjustments for speaker-listener perspective differences require resources.

So far, we have discussed automaticity in terms of efficiency. What about the other horsemen? Brothers et al. (2017) investigated to what extent prediction generation could be modulated by people's strategies (hence tapping into intentionality in the four horsemen). In their ERP study (Experiment 1), participants read medium-cloze (cloze = 50.7%) or low-cloze (0.7%) sentences, and they were instructed to simply comprehend the sentences or to try to predict the final word of each sentence. They found an N400 effect followed by a frontal positivity for low-cloze words relative to medium-cloze words, replicating previous findings (Federmeier & Kutas, 1999; Van Petten & Luka, 2012). Critically, these effects were larger when participants were instructed to predict than not, suggesting that participants' strategies can modulate effects of sentence constraints.

Brothers et al. (2017) additionally tested whether reliability of probabilistic cues affect prediction in a self-paced reading task (Experiment 2). The proportion of trials where a highly predictive sentence context was followed by the predictable word was varied across participants. Participants read target words faster when they followed a predictive context versus a non-predictive context. This predictability effect depended on the reliability of the cues; participants who read most of the predictive contexts followed by an unexpected word did not show this effect, and this effect was largest in participants who read most of the predictive contexts followed by the expected word (see also Hopp, 2016, for consistent findings in a visual world study). This suggests that people can inhibit prediction. The findings in Brothers et al. (2017) together demonstrate that people's strategies affect the degree of prediction, and prediction is not automatic in terms of intentionality or controllability.

In contrast to prediction-by-production, prediction-by-association seems largely automatic in terms of efficiency in L1 processing, because it is driven by lexical priming, and spreading of activation occurs rapidly in lexical priming (Neely, 1991; though later stages might be subject to strategic processing, Lau et al., 2013). Findings suggesting that prediction-by-association occurs rapidly comes from Kukona et al. (2011). In their Experiment 1, L1 English speakers listened to sentences with a semantically constraining verb such as *Toby arrests the crook*, while viewing images that corresponded to a typical patient of the constraining verb (the target: crook), a typical agent of the verb (the competitor: policeman), the mentioned subject (Toby), and unrelated distractors (surfer, florist). After hearing the verb (*arrests*), participants fixated not only the predictable target but also the unexpected competitor more than unrelated distractors. Participants started to fixate the target more than the competitor only after hearing the target. This finding suggests that predictive eye movements are (at least partly) driven by thematic priming (Ferretti et al., 2001) – priming from a verb to its typical agent and patient (irrespective of whether the agent/patient plausibly follows the verb). In Experiment 2, they used passive sentences such as *Toby was arrested by the policeman* and similarly found that participants fixated both the target (e.g., a policeman) and the competitor (e.g., a crook) over distractors after hearing the verb. But unlike in Experiment 1, they found that participants fixated the target more than the competitor before hearing the target. This finding suggests that the predictive eye movements are also driven by contextual information (i.e., not only by thematic priming), but context-based prediction may take more time to be generated. Crucially, thematic priming contributed to the early stage of predictive eye movements, suggesting that association-based prediction occurs rapidly.

Note that Gambi et al. (2016) used similar active sentences to those in Kukona et al.'s (2011) Experiment 1 and found no fixation bias towards an unexpected typical agent of a verb. After hearing an agent followed by a verb (e.g., *Pingu will ride*), L1 adults and children looked at appropriate patients (e.g., a horse) but not an inappropriate agent (e.g., a cowboy). As the authors discuss, one of the differences between the two studies was that the speech rate was much slower in Gambi et al. than in Kukona et al. Thus, it could be that participants in Gambi et al. had enough time to use the contextual information for prediction and predicted an appropriate referent by considering the syntactic structure (i.e., the upcoming referent is going to be a patient of the verb).

Prediction-by-association appears largely automatic in terms of controllability because people cannot stop the process. However, there is also evidence that people's strategies affect the degree of semantic priming (Lau et al., 2013), suggesting that spreading of activation may not be fully automatic in terms of intentionality or controllability (at least at late stages, see the discussion above). Kukona et al. (2016) found that people with higher comprehension skills (measured in several different tests) were better at inhibiting implausible referents that shared features with predictable referents (e.g., showing fewer looks to a white car after hearing *The boy will eat the white*…). People with higher comprehension skills may be better at controlling activation of inappropriate information (see also Peters et al., 2018), suggesting that prediction-by-association may not be fully automatic in terms of controllability or intentionality.

Automaticity of prediction in L2

Language comprehension is generally more resource-demanding (i.e., less automatic in terms of efficiency) in L2 than in L1 (Segalowitz & Hulstijn, 2009). L2 speakers may be slower or less likely to predict than L1 speakers because it takes more time and resources for L2 speakers to access lexical representations and build up sentence meaning that they can use for prediction. Language production in L2 also appears to be non-automatic in terms of efficiency because L2 production (picture naming) requires more time than L1 production (Ivanova & Costa, 2008), and lexical selection and monitoring processing are affected by task demands in L2 production (Declerck & Kormos, 2012). Consequently, prediction-by-production is likely to be less automatic in L2 than in L1, and L2 speakers may not predict as efficiently as L1 speakers. In addition, prediction-by-association in L2 may be less efficient than that in L1, because spreading of activation may be slower and lexical links may be weaker in L2 than in L1. In line with this, Dijkgraaf et al. (2019) found that predictive eye movements based on semantic associations were slower and weaker in L2 than in L1. In sum, non-automaticity involved in comprehension and production are both likely to account for reduced prediction in L2 speakers.

We discuss why prediction in L2 is limited compared to prediction in L1 in relation to (non-)automaticity in L2 processing. The first possibility is that similarity between L1 and L2 modulates prediction in L2 (Dussias et al., 2013; Foucart et al., 2014). Processing of lexical or grammatical features that are absent in L1 may be particularly subject to resource limitations because rules or representations in L1 (which L2 speakers can arguably apply more automatically) cannot be applied. Thus, L2 speakers may have difficulty in using L2-specific linguistic features for prediction (Bobb et al., 2015; Foucart & Frenck-Mestre, 2011). Another possibility is that L2 speakers may be slower to pre-activate L2 words because of interference from lexical representations in their L1. Lexical representations in L1 may compete when L2 speakers access L2 words in comprehension (Spivey & Marian, 1999) and production (Costa et al., 2006; Costa & Santesteban, 2004). For example, L2 speakers may try to pre-activate phonological representations in both L1 and L2 even when comprehending sentences in L2, and this may delay access to phonological representations in L2 and make prediction of phonology less efficient. However, there are other factors that may account for L1-L2 differences in the degree of prediction, some of which are unrelated to automaticity such as differences in frequency biases. For example, L1 and L2 speakers are likely to differ in how often they are exposed to particular combination of words, so they may predict different words (see Kaan, 2014). In this review, we discuss accounts that are related to automaticity (cf. Hopp, 2010, 2018).

Several studies have shown L2 speakers' difficulty in making predictions by using linguistic features that do not occur in their L1. Lew-Williams and Fernald (2010) investigated whether L2 Spanish speakers (with L1 English) use grammatical gender cues efficiently to identify target objects like L1 speakers in the visual world paradigm. Participants listened to sentences such as *Encuentra la pelota*. ('Find the$_{\text{-fem}}$ ball$_{\text{-fem}}$.'), while viewing two objects that had either the same or different gender in Spanish. L1 speakers fixated the target object more quickly when two objects had a different gender (i.e., when the article gender matched only the target object) compared to when they had the same gender, suggesting that L1 speakers used the grammatical gender cues efficiently to identify the target object. In contrast, L2 speakers did not benefit from the gender cues; they were no quicker to fixate the target object when the article gender allowed them to narrow down possible referents to a single object than when it did not. Thus, L2 speakers may not use gender (which is absent in their L1) to facilitate processing of upcoming words to the same extent as L1 speakers do.

Using German, Hopp (2013) found that L2 German speakers (with L1 English) predicted upcoming objects based on a gender cue, but only L2 participants who produced gender correctly 97% of the time in another production task predicted as quickly as L1 speakers. These participants were near-native German speakers who had been exposed to German for 42 years on average. L2 participants who produced gender correctly 77% of the time showed delayed predictive eye movements, and these participants had been exposed to German for 24 years on average. Hence, extremely high L2 proficiency may be needed for L2 speakers to predict like L1 speakers (as suggested by the performance in the production task). In a follow-up study, Hopp (2016) showed that L2 speakers came to use gender for prediction after being trained on gender assignment. The delayed prediction in less proficient L2 speakers in Hopp (2013) or the lack of prediction in L2 speakers in Lew-Williams and Fernald (2010) may be explained by L2 speakers' difficulty in encoding grammatical gender (as the production task results suggest), or their difficulty in using gender efficiently, or both.

Mitsugi and MacWhinney (2016) investigated prediction using case markers in Japanese (cf. Kamide et al., 2003; Experiment 3) in L1 Japanese speakers and L2 Japanese speakers (with L1 English) who had good knowledge of Japanese case markers in an offline test (with 86.5% accuracy). Participants listened to Japanese sentences with either a dative structure (e.g., *Gakkou-de majimena*$_{\text{-Adj}}$ *gakusei-ga*$_{\text{-Nom}}$ *kibishii*$_{\text{-Adj}}$ *sensei-ni*$_{\text{-Dat}}$ *shizukani*$_{\text{-Adv}}$ *tesuto-o*$_{\text{-Acc}}$ *watashita*$_{\text{-V}}$.; 'At the school, the serious student quietly handed over the exam to the strict teacher.') or an accusative structure (e.g., *Gakkou-de majimena*$_{\text{-Adj}}$ *gakusei-ga*$_{\text{-Nom}}$ *kibishii*$_{\text{-Adj}}$ *sensi-o*$_{\text{-Acc}}$ *shizukani*$_{\text{-Adv}}$ *karakatta*$_{\text{-V}}$.; 'At the school, the serious student quietly teased the strict teacher.'). A dative NP is almost always followed by an accusative NP, so the occurrence of a theme is predictable (Kamide et al., 2003). L1 speakers showed increased looks to a theme (e.g., a test) in the dative structure relative to the accusative structure before the theme was mentioned, suggesting that they predicted the occurrence of a theme. In contrast, L2 speakers' eye movements were not affected by the case marker, and they directed their looks to the theme only after it was mentioned. These results also suggest that L2 speakers may not use linguistic cues that are absent in their L1 for prediction as efficiently as L1 speakers.

The more direct evidence for the role of L1-L2 similarity comes from studies showing that L2 speakers whose L1 has grammatical gender are more likely to use gender cues for prediction than L2 speakers whose L1 does not have grammatical gender. Dussias et al. (2013) tested the use of gender cues in L1 Spanish speakers and L2 Spanish speakers with L1 English or L1 Italian. Replicating Lew-Williams and Fernald (2010), they found that L1 speakers used gender cues to quickly identify the upcoming noun. L1 Italian speakers used the feminine article but not the masculine article (for which there seems no clear explanation). L1 English speakers with a similar L2 proficiency did not use articles, but those with a higher proficiency (measured in grammar and vocabulary tests) showed predictive eye movements. Their findings suggest that both L1-L2 similarity and L2 proficiency modulate prediction.

Foucart et al. (2014) investigated prediction during reading in L1 Spanish speakers and L2 Spanish speakers (with L1 French) using Spanish highly predictable sentence contexts (e.g., *El pirata tenía el mapa secreto, pero nunca encontró* …; 'The pirate had the secret map, but he never found…') which were followed by either the expected continuation (e.g., *el tesoro*$_{\text{-masc}}$; 'the treasure') or an unexpected but plausible continuation (e.g., *la gruta*$_{\text{-fem}}$; 'the cave'). If participants predict the expected noun, the unexpected article was incompatible with the noun, so it was expected to elicit a different ERP than the expected article (cf. Wicha et al., 2004). In this study, both L1 and L2 speakers showed more negative ERPs for unexpected articles than for expected articles. This effect did not differ across groups, suggesting that L2 speakers predicted to a similar extent as L1 speakers. The negativity for unexpected versus expected articles was also found during listening comprehension in both L1 Spanish speakers (Foucart et al., 2015) and L2 Spanish speakers (with L1 French) (Foucart et al., 2016). Like Dussias et al. (2013), Foucart et al. (2016) argued that language similarity between L1 and L2 may mediate prediction (see also Foucart, this volume, for more detailed discussion).

Ito et al. (2021) examined prediction based on verb constraints that are specific to the L2 (Vietnamese) in L1 Vietnamese speakers and L2 Vietnamese speakers (with L1 German). For example, the Vietnamese verb *bê* means to carry a heavy object with two hands, while its German translation equivalent *tragen* means to carry (not necessarily a heavy object or with two hands), so the verb has a different mapping between Vietnamese and German. If L2 speakers suffer from interference from German verb constraints (i.e., L1 transfer effects), they might predict anything that one can carry when they hear *bê*. Participants listened to Vietnamese sentences (e.g., *Nam bê một chồng sách*; 'Nam carries a [classifier] book') which contained either a different-mapping (e.g., *bê – tragen*) or a similar-mapping verb (e.g., *ăn – essen*; both mean 'to eat') between Vietnamese and German. Upon hearing the verb, both L1 and L2 speakers fixated objects that met the verb constraints. But critically, L2 speakers also fixated objects that did not meet the Vietnamese verb constraints but met the verb constraints of the German translation equivalent (e.g., light things that can be carried). Thus, L2 speakers may have difficulty using L2 specific lexical features for prediction (see also Van Bergen & Flecken, 2017) and this may explain the lesser degree of prediction in L2 speakers.

It is important to consider whether L2 speakers' difficulty using L2-specific lexical or grammatical features for prediction comes from non-automaticity (e.g., inefficiency) of processing these features or incomplete acquisition/different representations of these features. The former possibility assumes that L2 speakers know lexical representations or grammatical rules in L2 but cannot use them in an automatic manner like L1 speakers, which could be because they often learned these features via explicit learning (e.g., by memorizing rules) (Hulstijn, 2002). As application of the grammatical rules is not as automatized as in L1 speakers, L2 speakers may not be able to use these features before predictable input comes up. The latter possibility assumes that the information that L2 speakers can use is different from what L1 speakers use. For example, if L2 speakers of Spanish do not have a lexical representation that the noun *ball* is feminine in Spanish, they cannot determine that it is grammatical after the feminine article *la*. As we reviewed above, there are cases where L2 speakers do not predict using the grammatical rules they know fairly well, as an offline test suggests (Mitsugi & MacWhinney, 2016). Nevertheless, L2 speakers seem more likely to predict using L2 specific lexical/grammatical features as their proficiency increases (Dussias et al., 2013; Hopp, 2013, though effects of proficiency may not be robust at all levels, cf. Hopp, 2015). Thus, automaticity in using L2 specific lexical/grammatical features online seems to be enhanced as L2 proficiency increases.

L2 speakers seem to predict less than L1 speakers even when L2 specific lexical/grammatical features are not primary cues for prediction. In an ERP study, DeLong et al. (2005) examined context-based prediction by presenting L1 English speakers with a predictive context (e.g., *The day was breezy so the boy went outside to fly …*) followed by either the expected continuation (e.g., *a kite*) or an unexpected but plausible continuation (e.g., *an airplane*). The expected word began with a consonant and the unexpected word began with a vowel, or vice versa. If people predict the expected word (e.g., *kite*), encountering an unexpected article (e.g., *an*) was expected to interfere with processing. They found that the N400 amplitude at the article was smaller (more positive) when the cloze probability of the article was higher. Because this effect was found before the presentation of the noun, it was taken as evidence for prediction of the expected noun. It is important to note that the *a*/*an* article effect has not replicated in later studies (Ito et al., 2017b; Nieuwland et al., 2018). Notably, Nieuwland et al. (2018) conducted a large-scale replication ($N = 334$) of DeLong et al. (2005) and did not find any effect at the article (but see Nicenboim et al., 2020 for a meta-analysis showing a small article effect).

Using this design, Martin et al. (2013) examined context-based prediction in L1 English speakers and L2 English speakers (with L1 Spanish). L1 speakers showed a negativity for the unexpected versus expected article 250 – 400 ms after the article onset. However, L2 speakers showed no effect at the article, though they were familiar with the *a*/*an* rule and showed an N400 effect for the unexpected versus expected nouns like L1 speakers. However, given the replication failures discussed above, it is unclear to what extent the L1-L2 difference in this study demonstrates evidence for limited prediction in L2 speakers.

Ito, Pickering, et al. (2018) tested context-based prediction in L1 English speakers and L2 English speakers (with L1 Japanese) using the visual world paradigm. Participants listened to sentences containing a highly predictable word (e.g., *The tourists expected rain when the sun went behind the* <u>*cloud*</u>, …) while viewing four objects. The critical object represented the predictable word (e.g., *cloud*), an English competitor which shared initial phonemes with the predictable word (e.g., *clown*), or an unrelated word (e.g., *globe*). (We will discuss another condition later in this section.) L1 speakers fixated both target and English competitor objects over the unrelated object before the target word was mentioned, suggesting that they predicted the phonological form of the target word. L2 speakers fixated the target predictively but they fixated the English competitor only well after the target word was mentioned. Thus, L2 speakers predicted some information about the target word, but there was no evidence that they predicted the phonological information. Since the target words were predictable from the context, L2 speakers seem to predict less than L1 speakers even when L2 specific lexical or grammatical features are not a primary cue for prediction.

Similarly, Ito et al. (2017a) investigated prediction of word form and meaning in L2 English speakers (with L1 Spanish) using the design in Ito et al. (2016; see the previous section for details) and found no evidence for prediction in L2 speakers. They showed an N400 reduction for words that were semantically related to the predictable word relative to unrelated words. However, this effect was no stronger when the cloze probability was higher, which is inconsistent with a prediction account. Form-related words showed a similar N400 to unrelated words, and thus there was no evidence that L2 speakers predicted word form. L2 speakers showed a Late Positive Complex (LPC) effect for form-related words relative to unrelated words at a slower word presentation rate (700 ms per word). This effect was larger in sentences with higher than lower cloze values, suggesting that L2 speakers processed sentences incrementally, combining both top-down information (sentence meaning) and bottom-up information (form-similarity). Nevertheless, they did not predict like L1 speakers (i.e., there was no N400 reduction for form-related words). Together, these findings (Ito et al., 2017a; Ito, Pickering, et al., 2018; Martin et al., 2013) indicate that L2 speakers may not predict like L1 speakers even when predictions are not based on L2 specific knowledge.

We mentioned that another factor that could affect prediction in L2 speakers is interference from L1 (Costa et al., 2006; Spivey & Marian, 1999). Given that lexical representations in L1 may interfere even when the L1 is irrelevant (e.g., during L2 comprehension), such interference may occur largely automatically (e.g., in terms of controllability) and may interfere with prediction in L2 by delaying access to L2 words (Hopp, 2018). In the study discussed above, Ito, Pickering, et al. (2018) tested whether L2 English speakers (with L1 Japanese) predict phonological information of a highly predictable word in their L1 by including a Japanese competitor condition, in which the critical object name shared initial phonemes with the predictable word when translated into Japanese (e.g., competitor: *bear*; Japanese: *kuma*; target: *cloud*; Japanese: *kumo*). If co-activating the Japanese equivalent of the target word interfered with prediction, we would have expected more fixations on the Japanese competitor than the unrelated object, but we did not find such an effect. However, this study did not find any evidence for word form prediction in L2 speakers, so it is unclear whether L2 speakers predict form in L1 when they do so in L2. If they do, the lesser degree of prediction in L2 may arise not only from resource limitations but also from interference from the L1. Regarding this issue, there is evidence suggesting that top-down language control affects lexical co-activation, such that lexical access to L1 words during L2 comprehension is facilitated by increased exposure to the L1 (Hoversten & Traxler, 2020). An interesting question for future research is whether lexical activation in L1 occurs automatically when people comprehend predictive sentences in L2 (cf. Ito et al., 2021) and whether that interferes with prediction in L2.

This review of studies suggests that whether L2 speakers predict also depends on the type of information involved. In fact, many studies found evidence for prediction of semantic information (Chambers & Cooke, 2009; Chun & Kaan, 2019; Dijkgraaf et al., 2017; Ito, Corley, et al., 2018). Some studies found evidence for prediction using syntactic information (Foucart et al., 2014, 2016; Hopp, 2013) while others did not (Lew-Williams & Fernald, 2010; Mitsugi & MacWhinney, 2016). In an investigation of semantic and syntactic prediction, Hopp (2015) found that L2 speakers used semantic information but not syntactic information for prediction. Upon hearing a verb in SVO sentences like *Der*$_{\text{-Nom}}$ *Wolf tötet*$_{\text{-V}}$ *gleich den*$_{\text{-Acc}}$ *Hirsch* ('The wolf will soon kill the deer'), L2 German speakers (with L1 English) looked at an appropriate patient (e.g., a deer) before it was mentioned, suggesting that they used verb semantics to predict an upcoming referent. However, they also tended to look at the same image (e.g., a deer) in OVS sentences like *Den*$_{\text{-Acc}}$ *Wolf tötet*$_{\text{-V}}$ *gleich den*$_{\text{-Nom}}$ *Jäger*. ('The hunter will soon kill the wolf'), suggesting that they did not use the case marker to predict that the upcoming referent would be an agent (rather than a patient) of the verb. Grüter et al. (2020) investigated prediction based on classifiers in Chinese and found that L2 speakers (with varied L1) tended to be distracted by objects that were grammatically inappropriate but matched a semantic class of a classifier (e.g., classifier *tiao* – long and flexible). This finding suggests that L2 speakers may be more likely to use semantic information (a semantic class of a classifier) than syntactic information (grammatical match between a classifier and a noun) for prediction. However, the studies that investigated prediction of phonological or orthographic word form have never found evidence for prediction in L2 speakers (Ito et al., 2017a; Ito, Pickering, et al., 2018; Martin et al., 2013). Thus, L2 speakers often use semantic information for prediction, while they are less likely to use syntactic information, and may not use word form information.

This pattern can be explained by prediction-by-production models (Pickering & Gambi, 2018; Pickering & Garrod, 2013), because in production, people first activate the semantic information of the word they are going to produce (conceptualization stage), then the syntactic information (grammatical encoding stage), and finally the phonological information (phonological encoding stage) (e.g., Levelt et al., 1999). Prediction-by-production involves the same stages, but the later stages may not occur when there is not enough time or resources available (e.g., prediction stages may be "cut-off" after the syntactic stage if time does not allow people to proceed to the phonological stage), because each stage in production entails some non-automaticity (e.g., Ferreira & Pashler, 2002; Garrod & Pickering, 2007; Hartsuiker & Barkhuysen, 2006). Thus, under this account, prediction of semantic information is more likely to occur than prediction of syntactic information, and prediction of syntactic information is more likely to occur than prediction of word form. Prediction-by-production during L2 processing may be particularly subject to resource limitations because of fewer resources available during L2 processing compared to L1 processing, making it less likely for L2 speakers to complete the full stages of prediction-by-production.

In the domain of L2 prediction research, there is not much evidence that relates to intentionality or controllability in the four horsemen (i.e., to what extent L2 speakers can control predictive processing). Regarding awareness, Curcic et al. (2019) used a Fijian-based miniature language and tested whether L2 speakers used a gender-marking rule to predict an upcoming noun based on the gender of a determiner. They found that L2 speakers who were aware that determiners could be used for prediction were more likely to predict than those who were unaware. This finding suggests a role of awareness, and prediction in L2 is non-automatic in terms of awareness.

Regarding efficiency, Ito, Corley et al. (2018; Experiment 2) repeated the experiment investigating effects of working memory on predictive eye movements with L2 speakers of English (see the previous section for details) and found similar effects of working memory in L2 speakers. L2 speakers (with varied L1s) made predictive eye movements when there was no working memory task, but the concurrent working memory task delayed these eye movements. This finding suggests that working memory mediates predictive eye movements in L2 speakers and prediction is non-automatic in terms of efficiency – in line with the findings in L1 speakers. This finding is also consistent with Kaan (2014), who argued that the same factors (e.g., working memory) mediate prediction in L1 and L2.

Conclusion

We evaluated how automatic prediction is by reviewing studies on prediction in L1 and L2 speakers. A large body of evidence suggests that prediction requires time and resources, and it tends to be reduced in L2 speakers, who have fewer resources available during comprehension compared to L1 speakers. These findings suggest that prediction is non-automatic in terms of efficiency in the four horsemen. There is also evidence suggesting that people can control predictive processing, at least in L1, suggesting that prediction is non-automatic also in terms of intentionality and controllability. We suggest that prediction-by-association, which complements prediction-by-production, occurs rapidly and is largely automatic in terms of efficiency. However, evidence for such prediction is rather scarce, and further investigation will be required to separately assess how automatic prediction-by-association is in both L1 and L2 comprehension. Our review also suggests that L2 speakers often predict semantic information like L1 speakers in simple sentences and can sometimes predict syntactic information but may be slower to predict or may not predict word form information. The available evidence is in line with the production-based prediction accounts: Because production requires time and resources, the later stages of prediction-by-production (syntactic encoding and phonological encoding) may not occur when there are fewer cognitive resources available, as is the case in L2 comprehension.

References

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer , Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes; Applications* (pp. 1–40). Lawrence Erlbaum Associates.

Bobb, S. C., Kroll, J. F., & Jackson, C. N. (2015). Lexical constraints in second language learning: Evidence on grammatical gender in German. *Bilingualism: Language and Cognition*, 18(3), 502–523. https://doi.org/10.1017/S1366728914000534

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89(1), 1–47. https://doi.org/10.1037/0033-295X.89.1.1

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, 93, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35(4), 1029–1040. https://doi.org/10.1037/a0015901

Chun, E., & Kaan, E. (2019). L2 prediction during complex sentence processing. *Journal of Cultural Cognitive Science*, 3(2), 203–216. https://doi.org/10.1007/s41809-019-00038-0

Cook, A. E., & Meyer, A. S. (2008). Capacity demands of phoneme selection in word production: New evidence from dual-task experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 886–899. https://doi.org/10.1037/0278-7393.34.4.886

Corps, R. E., Brooke, C., & Pickering, M. J. (2021). Do comprehenders predict from their own perspective? Manuscript in preparation

Costa, A., Heij, W., & La Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, 9(2), 137. https://doi.org/10.1017/S1366728906002495

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, 69, 42–71. https://doi.org/10.1111/lang.12321

Declerck, M., & Kormos, J. (2012). The effect of dual task demands and proficiency on second language speech production. *Bilingualism: Language and Cognition*, 15(4), 782–796. https://doi.org/10.1017/S1366728911000629

Dell, G. S. (1986). A spreading-activation theory of retreival in sentence production. *Psychological Review*, 93(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915

Fayol, M., Largy, P., & Lemaire, P. (1994). Cognitive overload and orthographic errors: When cognitive overload enhances subject–verb agreement errors. A study in French written language. *The Quarterly Journal of Experimental Psychology Section A*, 47(2), 437–464. https://doi.org/10.1080/14640749408401119

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41, 469–495. https://doi.org/10.1006/jmla.1999.2660

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 28(6), 1187–1199. https://doi.org/10.1037/0278-7393.28.6.1187

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 44(4), 516–547. https://doi.org/10.1006/jmla.2000.2728

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, 14(3), 379–399. https://doi.org/10.1017/S136672891000012X

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1461–1469. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2015). How do you know I was about to say "book"? Anticipation processes affect speech processing and lexical recognition. *Language, Cognition and Neuroscience*, 30(6), 768–780. https://doi.org/10.1080/23273798.2015.1016047

Foucart, A., Ruiz-Tada, E., & Costa, A. (2016). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 19(1), 213–219. https://doi.org/10.1017/S1366728915000486

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. *Cognition*, 157, 340–351. https://doi.org/10.1016/j.cognition.2016.10.003

Garrod, S., & Pickering, M. J. (2007). Automaticity of language production in monologue and dialogue. In A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (pp. 1–20). Psychology Press. https://doi.org/10.4324/9780203968512-7

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, 43(1), 21–30. https://doi.org/10.1111/j.1469-8986.2006.00383.x

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, 35(2), 221–234. https://doi.org/10.1080/23273798.2019.1648840

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, 21(1–3), 181–204. https://doi.org/10.1080/01690960400002117

Hartsuiker, R. J., & Moors, A. (2017). On the automaticity of language processing. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 201–225). American Psychological Association. https://doi.org/10.1037/15969-010

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120(4), 901–931. https://doi.org/10.1016/j.lingua.2009.06.004

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014

Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32(2), 277–307. https://doi.org/10.1177/0267658315624960

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, 17, 5–27. https://doi.org/10.11431/secondlanguage.17.0_5

Hoversten, L. J., & Traxler, M. J. (2020). Zooming in on zooming out: Partial selectivity and dynamic tuning of bilingual language control during reading. *Cognition*, 195, 104118. https://doi.org/10.1016/j.cognition.2019.104118

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research*, 1706, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18(3), 193–223. https://doi.org/10.1191/0267658302sr207oa

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, 86, 157–171. https://doi.org/10.1016/j.jml.2015.10.007

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017a). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(4), 635–652. https://doi.org/10.1037/xlm0000315

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017b). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience*, 32(8), 954–965. https://doi.org/10.1080/23273798.2016.1242761

Ito, A., Nguyen, H. T. T., & Knoeferle, P. (2021). Anticipatory use of verb and classifier constraints in Vietnamese-German bilinguals. Manuscript in preparation

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, 98, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 127(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa
Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8
Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., Pugh, K. R., Shankweiler, D. P., & Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica*, 171, 72–84. https://doi.org/10.1016/j.actpsy.2016.09.009
Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, 119(1), 23–42. https://doi.org/10.1016/j.cognition.2010.12.002
Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, 61(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004
Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, 25(3), 484–502. https://doi.org/10.1162/jocn_a_00328
Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. The MIT Press.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–75. https://doi.org/10.1017/S0140525X99001776
Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003
Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001
Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881
Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology*, 67(1), 263–287. https://doi.org/10.1146/annurev-psych-122414-033550
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Lawrence Erlbaum Associates.
Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, 142, 107427. https://doi.org/10.1016/j.neuropsychologia.2020.107427
Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. a, Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, 7, e33468. https://doi.org/10.7554/eLife.33468
Peters, R. E., Grüter, T., & Borovsky, A. (2018). Vocabulary size and native speaker self-identification influence flexibility in linguistic prediction among adult bilinguals. *Applied Psycholinguistics*, 39(6), 1439–1469. https://doi.org/10.1017/S0142716418000383

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3), 105–110. https://doi.org/10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329–392. https://doi.org/10.1017/S0140525X12001495

Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review*, 88(5), 454–462. https://doi.org/10.1037/0033-295X.88.5.454

Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in Psychology*, 2, 1–14. https://doi.org/10.3389/fpsyg.2011.00307

Segalowitz, N., & Hulstijn, J. H. (2009). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). Oxford University Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing, II: Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, 10(3), 281–284. https://doi.org/10.1111/1467-9280.00151

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, 1(6), 209–216. https://doi.org/10.1016/S1364-6613(97)01070-X

CHAPTER 3

Second language prediction ability across different linguistic domains

Evidence from German

Judith Schlenter and Claudia Felser
University of Potsdam

Previous research suggests that second language (L2) speakers can reliably use lexical-semantic cues for predictive processing but have difficulty using morphosyntactic ones. Here we report the results from a visual-world eye-tracking study that tested first language (L1) Russian/L2 German speakers in two parallel experiments, asking whether morphosyntactic prediction is indeed more limited in an L2 compared to semantic prediction. In Experiment 1 both the L2 speakers and L1 German-speaking controls showed evidence of using selectional restriction information to anticipate an upcoming direct object. Unlike what was reported previously, in Experiment 2 the same group of L2 speakers also showed evidence of using morphological case information predictively. However, between-group differences suggested that L2 speakers had more difficulty integrating competing cues during processing than L1 speakers.

Introduction

Non-native or second language (L2) speakers have been reported not to be able to use all available sources of linguistic information to the same extent as native or first language (L1) speakers do for prediction. Previous findings indicate that lexical-semantic information, for example, can be used as a predictive cue in L2 processing (e.g., Chambers & Cooke, 2009; Dijkgraaf et al., 2017, 2019; Ito, Corley, & Pickering, 2018), whereas morphosyntactic information is used less consistently (e.g., Dussias et al., 2013; Hopp & Lemmerth, 2018; Hopp, 2015; Mitsugi & MacWhinney, 2016; Frenck-Mestre et al., 2019). The investigation of prediction as a fast-operating mechanism has the potential to shed light on some unresolved questions regarding L1 and L2 differences in real-time processing, including the question of whether L1-L2 differences are mostly restricted to certain linguistic domains. Several L2 processing theories assume that the processing of (morpho-)syntactic information

https://doi.org/10.1075/bpa.12.03sch

is particularly vulnerable in L2 speakers, either because the mapping of lexical to syntactic information is a 'bottleneck' for them (Hopp, 2018) or because they put less weight on grammatical versus other types of information during processing than L1 speakers (Clahsen & Felser, 2006, 2018; see also Cunnings, 2017). However, the results from previous studies examining prediction in an L2 are not easily comparable because of differences between participants, experimental designs and the language combinations tested.

We will take 'prediction' to refer to the "pre-activation/retrieval of linguistic input before it is encountered by the language comprehender" (Huettig, 2015, p. 122) and may use this term interchangeably with 'anticipation.' Following a brief literature review, we present the results from a visual-world eye-tracking study that controls for the factors mentioned above by testing the same participants (proficient Russian L1/German L2 speakers) in two parallel experiments. We ask whether morphosyntactic prediction is indeed more limited in an L2 than in an L1, and more limited in comparison to semantic prediction. The chapter closes with a discussion of possible reasons for why our results differ from previous findings regarding L2 speakers' ability to use morphosyntactic information predictively, and of their theoretical implications.

Semantic prediction

L2 speakers have repeatedly been shown to be able to use lexical-semantic information (together with their world knowledge) to predict upcoming linguistic information. In visual-world eye-tracking studies L2 speakers used the lexical semantics of verbs to predict an upcoming patient or theme argument (L2 French: Chambers & Cooke, 2009; L2 English: Dijkgraaf et al., 2017; Ito, Corley, & Pickering, 2018). In Example (1) from Ito, Corley, and Pickering (2018), participants directed their gaze towards the depicted target object (here, a scarf) upon listening to the verb in sentence (1b), that is, before this object was mentioned.

(1) a. The lady will find the scarf.
b. The lady will fold the scarf.

In their experimental design, a baseline condition (1a) was compared to a condition in which the verb restricts the selection of upcoming arguments, here to objects that are foldable (1b).

Although L2 speakers have been shown to use lexical-semantic information predictively, their semantic prediction ability appears to be modulated by factors such as L2 proficiency (e.g., Chambers & Cooke, 2009; Peters et al., 2015). L2 semantic prediction may not always be as efficient as L1 semantic prediction (Ito, Pickering, & Corley, 2018; Dijkgraaf et al., 2019). For example, Ito, Pickering, and Corley (2018), who tested the prediction of a specific lexical item (and phonological form) based on the prior sentence context, report a difference between L1 and L2 speakers in the time course of prediction; for an extensive discussion on semantic prediction in L1 and L2 processing, see Dijkgraaf et al. (2019). Additionally, L2 semantic prediction may depend on the experimental design and materials, such as whether simple or more complex sentences are presented (Dijkgraaf et al., 2019; see also Chun et al., this volume), as well as on the participant groups tested.

The results from an eye-tracking study conducted by van Bergen and Flecken (2017) show that L2 semantic prediction can be affected by the existence of a matching predictive cue in the L1. The authors tested whether advanced L2 learners of Dutch with German, English or French as their L1 could predict the positioning of an object after encountering a placement verb. Whereas German, like Dutch, specifies the position of the object of placement verbs (Dutch: *zetten*, German: *stellen* – 'put.STAND' and Dutch: *leggen*, German: *legen* – 'put.LIE'), this is not specified in English or French, which use a general placement verb (English: *put*, French: *mettre*) only. Whereas both L1 Dutch and L1 German speakers demonstrated anticipation of the object's perceptual features in Dutch, that is, whether it was standing or lying, L1 English and L1 French speakers did not.

To summarize, proficient L2 speakers can use lexical-semantic information predictively in the L2, possibly aided by familiarity with the relevant predictive cues from their L1, although they may not always use semantic information as efficiently as L1 speakers do.

Morphosyntactic prediction

Morphosyntactic prediction in L2 processing has been reported to be limited or could not be attested. In the following, we will focus on the predictive use of morphological case; for a discussion on predictive gender agreement in L2 processing, see Hopp (2018).

Previous visual-world eye-tracking studies found that L1 but not L2 speakers anticipated an upcoming argument based on case marking on prior argument phrases. A study by Mitsugi and MacWhinney (2016) on Japanese, a verb-final language, examined whether participants were more likely to anticipate a theme argument after the canonical argument order Nominative > Dative and the scrambled order Dative > Nominative than after the order Nominative > Accusative. The latter served as a baseline condition because only a monotransitive verb can follow here, so that no additional argument should be expected. In neither of the two ditransitive verb conditions did intermediate-level L2 speakers with English as their L1 show anticipation of a theme argument. This could potentially be due to a lack of familiarity with using case cues for prediction among the L2 group. Note that English lacks a proper case marking system and marks case only on pronouns, and primarily relies on word order for determining thematic roles.

In a more recent study on Korean, Frenck-Mestre et al. (2019) tested two different L2 groups, L1 Kazakh and L1 French speakers. Unlike Korean and Kazakh, French has no morphological case marking and also has a less flexible word order than both Korean and Kazakh. The results indicated that L1-L2 overlap played a role in L2 processing: Although neither L2 group showed any anticipatory effect, the L1 Kazakh group showed a more nativelike pattern for later sentence interpretation than the L1 French group. During the processing of the sentence-final verb, L1 Kazakh speakers looked more frequently at the correct than at the incorrect visual scene, both of which were shown next to each other, not only for canonical but also for scrambled word order. In addition, the L1 Kazakh group was more accurate than the L1 French group in selecting the correct visual scene in the behavioral task. The accuracy and eye-movement data indicated that L1 French speakers prioritized word order over case. Since the L2 speakers had been studying Korean for three semesters only and thus were probably not advanced L2 speakers, the results leave open the question of whether, at higher proficiency levels, L2 speakers whose L1 shares properties with the L2 (case marking, free word order) can use case predictively.

In a study by Hopp (2015) on German, the predictive use of case marking was investigated via object topicalization as illustrated in Example (2).

(2) a. *Der Wolf tötet gleich den Hirsch.*
[The wolf]-NOM kills soon [the deer]-ACC
'The wolf will soon kill the deer.'
b. *Den Wolf tötet gleich der Jäger.*
[The wolf]-ACC kills soon [the hunter]-NOM
'The hunter will soon kill the wolf.'

In the canonical SVO condition (2a), L1 German speakers anticipated a plausible theme (*Hirsch* 'deer'), and in the non-canonical OVS condition (2b) they anticipated a plausible agent (*Jäger* 'hunter'). In contrast, and irrespective of their German proficiency level, L1 English speakers showed no difference between the two word-order conditions and always anticipated the plausible theme of the canonical condition, that is, *Hirsch* 'deer' in Example (2). This shows that whereas the L1 speakers integrated morphosyntactic and semantic information successfully, the L2 speakers relied on word order and verb semantics only (in line with the proposal that L2 speakers' semantic prediction is intact).

Again, one possible explanation for the absence of morphosyntactic prediction in the L2 group could be the lack of similarity between the L1 and the L2. Additionally, the experimental design could have affected the outcome: Firstly, object topicalizations are relatively infrequent in German (e.g., Bader & Häussler, 2010) and, without facilitating discourse information, represent marked structures. Secondly, in the example from Hopp (2015), case on the sentence-initial NP is only marked on the article, and nominative (*der Wolf*) and accusative (*den Wolf*) may be hard to discriminate acoustically.

To summarize, previous studies have found L2 speakers to be unable to use morphological case as a predictive cue. However, it is unclear whether these results can be generalized to L2 processing as previous studies did not include any highly proficient L2 speakers or L2 speakers who were familiar with morphological case from their L1.

The current study

We used the visual-world eye-tracking paradigm to compare semantic and morphosyntactic prediction in L1 and L2 speakers of German. The L2 group consisted of Russian L1 speakers who had started to learn German at or after the age of seven. Hence, most of our L2 speakers learned German in school or university. Similar to German, Russian is a case-marking language with flexible word order. L2 speakers' proficiency was assessed through the Goethe placement test (Goethe Institute, 2010), a 30-item multiple choice test. Additional offline tests examined L2 speakers' lexical knowledge (specifically, their knowledge of verb meaning), and their knowledge of German case marking and subject-verb agreement. Our study was part of a larger investigation comprising a total of three experiments, two of which will be reported here.

Experiment 1 tested the use of lexical-semantic information to predict the animacy of the direct object noun in transitive event descriptions such as *The woman feeds/irons the black cat/blouse*. The verb restricted the selection of upcoming nouns to the category animate (e.g., *feed*) or inanimate (e.g., *iron*), as was assessed by a cloze test. The meaning of the verbs did not differ between German and their Russian translation equivalents; all verbs denoted the same action in both languages. Based on previous findings on semantic prediction, we expected L2 speakers to show an effect of prediction, that is, more looks to the target as compared to a distractor object once the verb has been encountered.

Experiment 2 tested the use of morphological case marking. This was done by manipulating the argument order in double object constructions (Dative > Accusative vs. Accusative > Dative). To increase the salience of the case cue, and taking into account the fact that Russian has no article system but marks case on adjectives, we inserted an adjective between the article and noun of the first postverbal NP (e.g., *der*-DAT *blühenden*-DAT *Pflanze* vs. *die*-ACC *blühende*-ACC *Pflanze* – 'the flowering plant'). Dative marking indicates a recipient and accusative marking a theme argument, so that the case marking on the first postverbal NP should trigger an expectation for either a theme or a recipient. If L2 speakers, in line with previous findings, were unable to use case information predictively, they should show the same gaze pattern for both argument orders.

Method

Participants

The L1 group consisted of 28 German native speakers (23 female, mean age: 25.96), none of whom reported being early bilingual. Four additional participants were tested but their data subsequently excluded due to bad calibration and/or unstable eye-tracking.

In the L2 group, data from 25 Russian L1 speakers (22 female, mean age: 27.64), none of whom reported being early bilingual, were included. Five additional L2 participants were tested but later excluded: Two were excluded due to bad calibration and/or unstable eye-tracking. The others were excluded because of their low accuracy in the behavioral task of the eye-tracking experiment (one participant), their relatively low German proficiency (scoring below 21 in the Goethe test, one participant), and/or their poor performance in the offline questionnaire (one participant). The Goethe test scores ranged from 21 to 29 (mean: 25.32), placing the L2 participants at the upper B2 to C1/C2 level according to the Common European Framework of Reference for languages (CEFR). The L2 speakers' average age of onset of German was 13.08 years, ranging from seven to 24 years. Their length of residence in Germany and thus, their years of immersion ranged from zero, for a person who was only visiting Germany but had studied German in Moscow, to 21 years (mean: 7.92 years).

All participants had normal vision or, when necessary, wore glasses or lenses and reported no speech or hearing disorders. They were recruited from among the student community at Potsdam and Berlin, and participated in exchange of course credit or a monetary compensation. All participants gave informed written consent, and all procedures were in accordance with the Declaration of Helsinki.

Design and materials

An example item from Experiment 1 is shown in (3), for the animate-biasing (3a) and the inanimate-biasing verb condition (3b). Curly brackets indicate the critical window for an anticipatory effect.

(3) a. *Die Frau füttert {die schwarze} Katze.*
The woman feeds {the black} cat
b. *Die Frau bügelt {die schwarze} Bluse.*
The woman irons {the black} blouse

The items for Experiment 1 were selected from a cloze test conducted with a group of 35 speakers who did not participate any further in this study. This norming study served to ensure that all items in the eye-tracking task included a verb that had a clear bias towards an animate or an inanimate noun. Nouns with the highest cloze probability score were used as direct object nouns in the eye-tracking task. Except for one item, all target and distractor objects had different phonological onsets.

An example item from Experiment 2 is shown in (4), for the canonical Dative/Recipient > Accusative/Theme order (4a) and the non-canonical Accusative/Theme > Dative/Recipient order (4b).[1]

(4) a. *Der Gärtner gibt der blühenden {Pflanze eilig}*
The gardener gives the-DAT flowering-DAT {plant quickly}
frisches Wasser.
fresh water
'The gardener quickly gives fresh water to the flowering plant.'
b. *Der Gärtner gibt die blühende {Pflanze eilig}*
The gardener gives the-ACC flowering-ACC {plant quickly}
dem Postboten.
the postman
'The gardener quickly gives the flowering plant to the postman.'

The article and adjective of the first post-verbal argument mark it either as a dative or accusative object and thus, either as a recipient or a theme argument. All nouns used as the first post-verbal argument had either feminine or neuter gender, because

1. An acceptability rating task demonstrated that sentences with the order Accusative > Dative were rated as less acceptable by a group of 42 native German speakers than sentences with the order Dative > Accusative. Sentences with the non-canonical order still received high acceptability scores, however. The results from the acceptability rating task (Experiment 2) and of the cloze test (Experiment 1) are available at the Open Science Framework website at https://osf.io/vpsj5/, along with complete lists of our experimental items and other supplementary material.

dative and accusative marking for masculine gender is acoustically hard to discriminate in German and there is no difference on the adjective (*dem kleinen Jungen* vs. *den kleinen Jungen* – 'the small boy'). The adjectives used could not plausibly refer to the final argument because the two were not semantically or pragmatically compatible and/or were gender incongruent. The final argument was inanimate in the canonical condition and animate in the non-canonical condition, with only one exception. The animacy of the first post-verbal argument varied.

The sentences were recorded in a sound-attenuated room and spoken by a female native speaker of standard German at a normal speaking rate. Cross-splicing was applied using Praat (Boersma, 2001), so as to ensure that the critical window was identical in length and prosody between conditions. The picture set consisted of non-colored drawings taken from the MultiPic database (Duñabeitia et al., 2018) complemented by new drawings and/or drawings taken from free databases. All pictures of human beings had the same style throughout the experiment. The pictures had a size of 400 × 400 pixels. Figure 1 shows examples of the visual displays used in each experiment, which were kept the same across experimental conditions, for the example stimulus items in (3) and (4). In Experiment 1, the subject/agent was always displayed at the bottom center and the position of the target and distractor object was counterbalanced. In Experiment 2, the position of the four pictures was randomized.

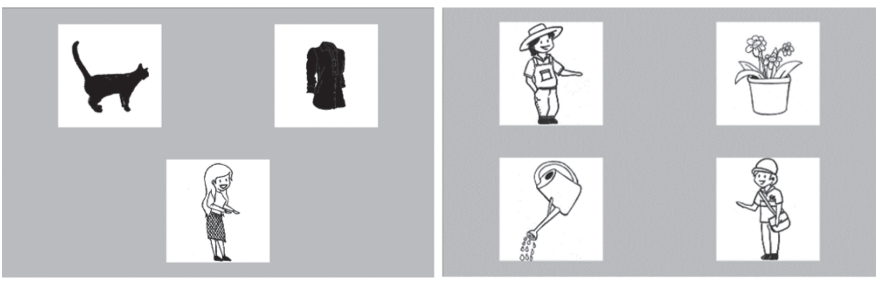

Figure 1. Visual displays for Experiment 1 (left) and Experiment 2 (right)

All experimental items appeared in two conditions that were distributed equally across two presentation lists using a Latin square design. In total, each participant encountered 108 sentences, 24 for Experiment 1 (12 in each condition), 28 for Experiment 2 (14 in each condition), 28 for another experimental set investigating subject-verb agreement, and 28 filler sentences. The order of presentation was pseudo-randomized, so there were no more than two items from the same experimental set in a sequence. All experiments were administered together in a single experimental session.

A forced-choice completion task, administered as an offline questionnaire, tested L2 speakers' knowledge of German case marking and subject-verb agreement. They were presented with slightly shortened versions of the materials from the eye-tracking experiment and had to indicate the grammatically correct sentence completion as shown in Example (5).

(5) *Der Gärtner gibt die blühende Pflanze…*
'The gardener gives the flowering plant…'
a. *frisches Wasser* ('fresh water')
b. *dem Postboten* ('the postman')

The offline questionnaire further included a list of the verbs the participants encountered in Experiment 1 and asked participants to indicate whether they were familiar with them.

Apparatus and experimental procedure

Eye-movements were tracked with an SMI RED eye-tracker at a sampling rate of 120 Hz. Participants sat at a distance of around 65 cm in front of the presentation screen, which had a resolution of 1680 × 1050 pixels. Viewing was binocular and unrestricted. A nine-point calibration procedure was used.

All participants received the same written instructions and each experimental session started with four practice trials, after which participants could ask any remaining questions they had about the experimental procedure. The participants listened to the spoken sentences via headphones. Each trial started with a one-second preview of the visual display. After each sentence, the visual display remained on the screen for another 800 milliseconds (ms). Half of the sentences in the eye-tracking experiment were followed by a written statement referring to the prior spoken sentence and requiring a true-false judgement. To respond with either 'yes' or 'no,' participants had to press the corresponding button on a gamepad. There was no time pressure and participants were encouraged to respond as accurately as possible. The number of correct 'yes' and 'no' responses was counterbalanced. Each experimental session contained two breaks, which allowed the participants to take a rest. Each break was followed by new gaze calibration/validation, which was repeated whenever necessary.

The L1 speakers only participated in the visual-world eye-tracking experiment, which took them around 30 minutes to complete. The L2 speakers additionally completed an offline questionnaire afterwards, which included the forced-choice completion task, testing L2 speakers' knowledge of case marking and subject-verb agreement, and a vocabulary list. For the L2 speakers, an experimental session lasted approximately 45 minutes.

Data analysis procedure

Verbs which L2 speakers had marked as unknown in the offline questionnaire were excluded from the analyses of Experiment 1 on a by-participant basis (24 trials in total). We excluded five experimental items from the analyses of Experiment 2 as they contained a potential ambiguity such that in the canonical word order condition (4a), the picture of the recipient in the non-canonical condition (4b) might also have been a plausible theme.[2] Additionally, two trials had to be removed from the L1 data set due to a participant coughing and to a sound problem, as well as one trial from the L2 data set that was accidentally skipped, plus two trials for which L2 speakers reported unknown vocabulary. The third block after the second break for one of the L1 speakers, and the first block up to the first break for one of the L2 speakers, were also excluded, in the first case because eye-tracking became unstable towards the end of testing, and in the second case because the participant still had some difficulty with the experimental task after the practice trials.

For data preparation we used the eyetrackingR package (Dink & Ferguson, 2015). The exact on- and offset information previously determined in Praat (Boersma, 2001) was used to define the onset of the critical window for an anticipatory effect. In Experiment 1 this was the offset of the verb and in Experiment 2 the offset of the adjective of the first postverbal argument. Onsets corrected by 200 ms to account for eye movement latency (Matin et al., 1993) were set to zero. The length of the critical window was determined based on the average length across items, so each item contributes a similar amount of data. As data points were aggregated into 50 ms time bins, we selected 500 ms for Experiment 1 (mean length: 515.5 ms, SD: 81.5 ms) and 950 ms for Experiment 2 (mean length: 967 ms, SD: 136 ms).

For statistical analyses, generalized linear mixed-effects models were computed in R (R Core Team, 2020), using the lme4 package (Bates et al., 2015) and, to obtain *p*-values, the lmerTest package (Kuznetsova et al., 2017). To take into account the trajectory of the effect (Mirman et al., 2008), orthogonal time polynomials were included in the models. The best-fitting model was selected based on the lowest AIC value of the converging models (Matuschek et al., 2017) with the random effect structure varying from minimal to maximal (Barr et al., 2013). Instead of raw binary values, we used the counts of events. By aggregating the binary values of a time bin into a count one can reduce the size of the dataset and, at the same time, reduce the correlation between neighboring observations, addressing

2. For example, for the sentence fragment *The student presents to his current girlfriend…* the intended theme argument was the inanimate object *motorbike*, but the simultaneously presented picture of *parents* would also possibly fit.

the problem of eye-movement-based dependencies as described by Barr (2008).[3] If not stated otherwise, we used deviation coding (0.5/−0.5) for the independent variables condition and group, both of which are categorical variables with two levels each. With the current coding scheme, the intercept shows the grand mean, that is, the estimate across conditions and groups for the average analysis window, while the effects of condition and group show the average difference between the two conditions and groups respectively (as for treatment coding). If a model based on this coding scheme revealed a significant difference for condition or group, we used treatment contrasts and relevelled our variables to obtain the estimate for each level. The specific formula can be found in the scripts and documentation provided at https://osf.io/vpsj5/.

Results

Truth value judgement and questionnaire data

The L1 group showed an overall response accuracy of 97% in the truth value judgement task (SD: 4%, range: 80–100%). The L2 speakers correctly responded to the statements with a mean of 93% (SD: 5%, range: 80–100%). Only one L2 participant was excluded because of less than 80% accuracy, which was set as a threshold for the inclusion in the eye-tracking analyses for all offline measures. The results thus indicate that participants paid attention to the stimulus items and had no problems in understanding their content.

For the forced-choice completion task that was part of the L2 offline questionnaire, accuracy scores ranged from 82% to 100% (mean: 97%, SD: 5%) for the case-marking set, showing that all L2 speakers whose data were included in the eye-tracking analyses for Experiment 2 had good knowledge of German case marking. As indicated in the participant section, one additional participant from the L2 group was excluded due to less than 80% accuracy in this task.

Eye-tracking data

In Experiment 1 we asked whether L1 and L2 speakers would look at the respective target object before encountering the direct object noun, thus demonstrating anticipatory eye-movements driven by the verb. Thus, fixations on the animate object

3. For the formula we used cbind(SamplesInAOI, (SamplesTotal – SamplesInAOI)) to analyze fixations of the target and competitor within each 50 ms time bin of every trial. The analyses differ from previous analyses reported in Schlenter (2019). The current analyses better account for the binomial distribution of eye-tracking data (e.g., Donnelly and Verkuilen, 2017). Since adding random slopes often led to non-convergence in the logistic regression, formulas differ from those in previous analyses. The current results largely replicate previous results.

after an animate-biasing verb were coded as 'target' looks and fixations on the inanimate distractor object as 'competitor' looks, and vice versa for inanimate-biasing verbs. Looks to the picture of the subject/agent were excluded. Figure 2 shows the proportion of looks to the target object relative to looks to target and competitor for the two language groups.[4]

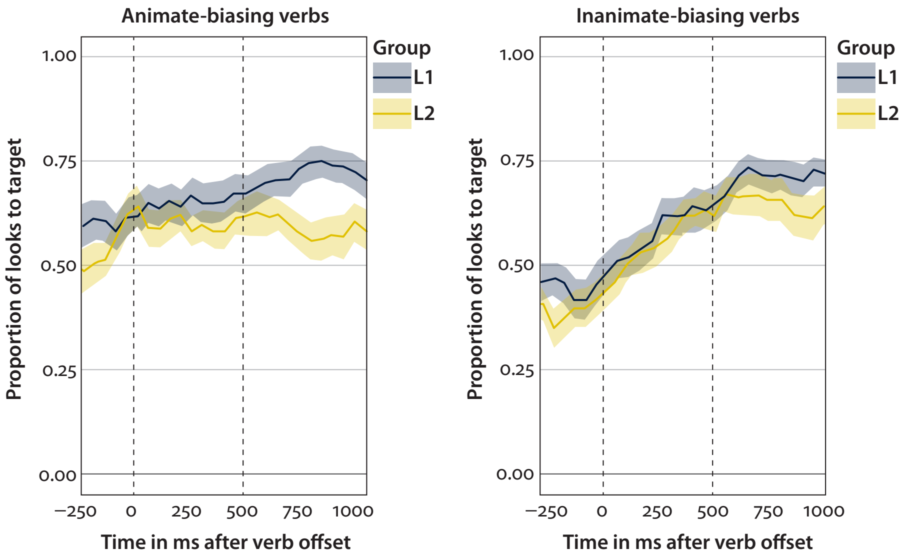

Figure 2. Proportion of looks to the target object relative to looks to target and competitor for the animate-biasing verbs (left) and inanimate-biasing verbs (right). Error bands represent standard errors. The dashed vertical lines indicate the critical time window.

The statistical model included the factors verb condition (animate-bias, inanimate-bias) and language group (L1, L2) as fixed effects. Because we expected that looks to the target would increase during the critical time window, the linear orthogonal time polynomial (ot1) was included as a fixed effect, as well as the interactions between factors. The random effects structure comprised subjects and items as random intercepts, the interaction between condition and linear time as by-item slope and the interaction between condition and linear time as by-subject slope. The model output is shown in Table 1.

The significant effect at the intercept in a positive direction indicates a target advantage across groups: Within the critical window, participants were more likely to fixate the target than the competitor. The effect of linear time indicates that this

4. For Experiment 1 and Experiment 2, additional graphs showing the proportion of looks for all pictures are available at https://osf.io/vpsj5/.

Table 1. Results of the logistic regression model for Experiment 1

	Est.	SE	*z*-value	*p*-value
Intercept (L1/L2)	0.581	0.178	3.258	0.001**
Condition	0.637	0.388	1.642	0.101
Group	0.515	0.276	1.867	0.062+
Linear Time	0.765	0.287	2.665	0.008**
Condition × Group	0.367	0.590	0.622	0.534
Condition × Linear Time	−1.297	0.617	−2.104	0.035*
Group × Linear Time	0.088	0.438	0.200	0.842
Condition × Group × Linear Time	0.200	0.984	0.203	0.839

target advantage increased over time. However, this increase was different for the verb conditions as signaled by the interaction between condition and linear time. The same model with the treatment-coded variable verb condition showed that for the animate-biasing verbs participants across groups showed a target advantage for the average time window, while the target advantage for the inanimate-biasing verbs developed over time (Linear Time, Est.: 1.413, SE: 0.432, *z*-value: 3.273, *p*-value: 0.001). This difference between verb conditions was probably enhanced by the preference for the animate entity in both groups, which is visible at verb offset in Figure 2. While the participants' gaze remained on the target for the animate-biasing verbs, they shifted their gaze to the target for the inanimate-biasing verbs, which contributes to the steeper slope for the latter (compare Barr et al., 2011).

Since the lack of significant between-group differences does not necessarily mean that both groups behaved in exactly the same way – in fact, the above model showed a marginal effect of group – we also computed the model with the treatment-coded variable group, so by re-levelling we received the estimate for each verb condition per language group. While the effect of linear time for the inanimate-biasing verbs was observed in both groups, the target advantage for the animate-biasing verbs, that is, a significant effect at the intercept, was only seen in the L1 group. As can be seen in Figure 2 for the animate-biasing verbs, the L2 group was less likely to fixate the target during the critical window in comparison to the L1 group. The results from these analyses thus point to slight differences, most notably for the animate-biasing verbs, although the pattern seen in the L2 group largely resembled that of the L1 group.

In the analysis for Experiment 2 we examined whether L1 and L2 speakers anticipate the final argument, a plausible theme or recipient argument, based on the case marking on the article and adjective of the first postverbal argument. Analyses included fixations on the theme (*Wasser* 'water') and recipient (*Postbote* 'postman'), excluding the two other pictures showing the agent (*Gärtner* 'gardener') and the first postverbal noun (*Pflanze* 'plant'). Recall that the critical window in Experiment 2 is the window where participants heard the first postverbal noun, and an anticipatory effect is expected to develop within this time window. The canonical argument order condition (Dative > Accusative) was taken as the baseline in the statistical analyses for Experiment 2. Figure 3 shows the proportion of looks to the accusative object/theme, that is, to the target for the canonical argument order relative to looks to target and competitor.

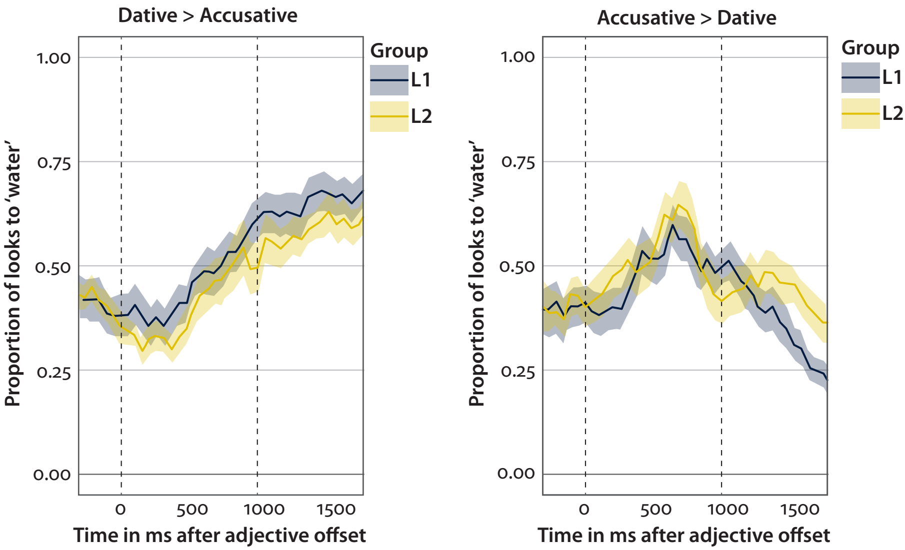

Figure 3. Proportion of looks to the theme (*Wasser* 'water') relative to looks to theme and recipient (*Postbote* 'postman') for the canonical order Dative > Accusative (left) and the non-canonical order Accusative > Dative (right). Error bands represent standard errors. The dashed vertical lines indicate the critical time window.

Here the factor argument order condition (canonical, non-canonical) was treatment-coded with the canonical condition as the reference level. If case was used predictively, participants should be less likely to fixate the theme in the non-canonical (4b) than in the canonical condition (4a). Since initial looks towards the competitor theme were expected for the non-canonical argument order, next to linear time, quadratic time was included as a second-order orthogonal polynomial (ot2) in our model to better capture the curvature in the data. The model formula included the factor argument order condition and group together with time (linear and quadratic) as fixed effects and their respective interactions. The best-fitting model that converged included subjects and items as random intercepts and the interaction between condition and group as by-item slope, as well as the interaction between condition and linear time as by-subject slope. The model output is shown in Table 2.

Table 2. Results of the logistic regression model for the critical time window in Experiment 2

	Est.	SE	*z*-value	*p*-value
Intercept (L1/L2, canonical)	−0.485	0.249	−1.947	0.052^{+}
Condition	0.354	0.271	1.307	0.191
Group	0.137	0.419	0.326	0.745
Linear Time	1.785	0.544	3.283	0.001^{**}
Quadratic Time	0.684	0.072	9.564	$< 0.001^{***}$
Condition × Group	−0.287	0.512	−0.561	0.575
Condition × Linear Time	−0.032	0.886	−0.036	0.971
Condition × Quadratic Time	−1.553	0.100	−15.497	$< 0.001^{***}$
Group × Linear Time	0.297	1.063	0.280	0.780
Group × Quadratic Time	0.097	0.143	0.676	0.499
Cond. × Group × Linear Time	0.547	1.733	0.316	0.752
Cond. × Group × Quadratic Time	0.529	0.200	2.644	0.008^{**}
Intercept (L1/L2, non-canonical)	−0.132	0.217	−0.606	0.545
Group	−0.150	0.361	−0.417	0.677
Linear Time	1.754	0.710	2.470	0.014^{*}
Quadratic Time	−0.868	0.070	−12.381	$< 0.001^{***}$
Group × Linear Time	0.845	1.304	0.648	0.517
Group × Quadratic Time	0.626	0.140	4.469	$< 0.001^{***}$

The model output shows no effect of argument order across the critical window (no effect on the intercept term), but an effect of linear and quadratic time, which indicate that theme fixations increased significantly for the canonical condition (4a) as predicted. In addition, the results show a significant interaction between argument condition and quadratic time. This interaction was further qualified by a three-way interaction between the factors condition, group and quadratic time. When taking the non-canonical condition (4b) as the reference level, the model shows an effect of linear time in a positive direction and of quadratic time in a negative direction as well as an interaction between group and quadratic time.

Both groups initially considered the theme argument, although implausible, in the non-canonical condition, however, as time further increased, participants were less likely to fixate the theme. Initial consideration of the competitor theme was slightly more pronounced in the L2 group, who then showed a steeper decrease in looking at the theme. Crucially, the condition by (quadratic) time interaction indicating a difference between argument orders over time was visible in both the L1 and the L2 group (L1, Est.: −1.288, SE: 0.140, z-value: −9.189, p-value: < 0.001; L2, Est.: −1.817, SE: 0.143, z-value: −12.706, p-value: < 0.001) as shown by the same model with the treatment-coded variable group and re-levelling. Hence, in the critical time window both groups largely patterned alike.

To examine whether the emerging anticipatory effect seen in the analysis of the critical window actually led to an overall effect of argument order condition in the post-critical time window, the 650 ms time window following the onset of the final argument (shifted 200 ms forwards) was also analyzed. The same model as for the critical window revealed an effect of group. Therefore, the model with the treatment-coded variable group was computed, the results of which are shown in Table 3, with the L1 group as reference level.

Table 3. Results of the logistic regression model for the post-critical time window in Experiment 2

	Est.	SE	*z*-value	*p*-value
Intercept (L1, canonical)	1.330	0.430	3.096	0.002**
Condition	−1.989	0.527	−3.777	< 0.001***
Group	−0.998	0.503	−1.982	0.047*
Linear Time	0.283	0.655	0.432	0.666
Quadratic Time	−0.135	0.102	−1.329	0.184
Condition × Group	1.471	0.667	2.207	0.027*
Condition × Linear Time	−1.698	0.905	−1.877	0.061+
Condition × Quadratic Time	−0.588	0.137	−4.294	< 0.001***
Group × Linear Time	−0.237	0.945	−0.251	0.802
Group × Quadratic Time	0.735	0.141	5.223	< 0.001***
Cond. × Group × Linear Time	1.319	1.312	1.005	0.315
Cond. × Group × Quadratic Time	0.211	0.194	1.086	0.277

As can be seen in Table 3, for the L1 group the emerging anticipation of the final argument in the critical window turned into an overall effect of condition upon its auditory presentation: Theme fixations were above chance for the canonical condition, and there was a significant difference between argument order conditions for the average post-critical window. Competitor (i.e., the theme argument) fixations continued to decrease as indicated by the condition by time interaction. For the L2 group, the preference for the theme over the recipient in the canonical condition (and vice versa for the non-canonical condition) further developed over time and did not result in an overall effect of condition. In Table 3, this is reflected in the group effect, and by the group by condition and group by time interactions.

Summary

Experiment 1 tested the predictive use of lexical-semantic information, with the level of prediction systematically controlled for by using verbs that semantically selected either animate or inanimate objects. Both L1 and L2 groups showed evidence of anticipating the upcoming direct object noun's animacy. Within a critical window for an anticipatory effect, participants across both groups were more likely to fixate the target than the competitor picture. The effect developed differently for the two verb conditions, probably resulting from a general preference for animate over inanimate entities not atypical in eye-tracking research (see e.g., Kamide et al., 2003, p. 139).

In Experiment 2 the same participants also anticipated the final argument while listening to a canonically ordered sentence fragment such as *Der Gärtner gibt der blühenden Pflanze eilig…* ('[The gardener]-NOM gives [the flowering plant]-DAT quickly…'). Looks to the picture of the target accusative object/theme ('water') increased across both groups while they were listening to the dative argument. Crucially, in both groups looks to the same picture ('water') developed differently for the non-canonical argument order (*Der Gärtner gibt die blühende Pflanze eilig…* '[The gardener]-NOM gives [the flowering plant]-ACC quickly…'). Although both groups initially considered the competitor theme, they became aware that a dative object/recipient (here, *dem Postboten* 'the postman') had to follow. Hence, Experiment 2 demonstrated that both L1 and L2 speakers were able to map the case marking on an argument to a thematic role in a timely enough manner so as to show a difference in the time course between the two argument order conditions (Dative > Accusative vs. Accusative > Dative). Evidence for this was already visible before the onset of the final argument, indicating that our participants were able to predict the upcoming thematic role.

Besides the above similarities between the L1 and the L2 group, in both experiments we also found some subtle differences between the two groups. In Experiment 1, the L2 group was less certain in their prediction than the L1 group, as seen for the animate-biasing verbs, supporting the claim that semantic prediction in L2 processing can be less efficient than in L1 processing (Dijkgraaf et al., 2019). Evidence for reduced certainty and for a delayed effect in L2 compared to L1 speakers were also seen in Experiment 2. The L2 group was more susceptible to competition between the theme and recipient arguments in the non-canonical condition than the L1 group, although this competition most clearly affected later information integration as seen in an analysis of the post-critical time window. This analysis showed that L2 speakers' integration of the final argument was delayed, while in the L1 group anticipation led to an immediate preference for the respective target.

Discussion

Besides confirming earlier findings showing that L2 speakers can use semantic information predictively (Experiment 1), the current study's most intriguing finding was the observed difference between the two argument order conditions in the L2 group in Experiment 2. This indicates that the L2 speakers we tested were sensitive to case information and were able to use this information predictively during real-time processing, a finding which contrasts with previous findings reported by Hopp (2015), Mitsugi and MacWhinney (2016) and Frenck-Mestre et al. (2019). There are several possible explanations for this discrepancy, which are not mutually exclusive. First, recall that we only included L2 speakers at an advanced to near-native level of proficiency who also demonstrated good knowledge of German case marking. Second, we chose an L2 group who had experience with case as a predictive cue. Our L2 group was comprised of native speakers of Russian, a language with an elaborate case system and flexible word order. Mitrofanova et al. (2019) showed that children below the age of six were already able to use case predictively in Russian. Unlike the participants in previous studies, our L2 participants might thus have benefited from familiarity with case as a predictive cue and L1-L2 similarity. Third, in our materials we tested the use of morphological case by manipulating argument order after ditransitive verbs (i.e., within the sentence's 'mid-field'). Although this resulted in a non-canonical word order (Nominative > Accusative > Dative), this order is probably less marked compared to the object topicalization structure (Accusative > Nominative) previously tested by Hopp (2015). Häussler and Bader (2011, p. 297) note that the markedness difference between the two alternative word orders in the mid-field, to the extent that such a difference exists, is rather small, which is in line with the results from our acceptability rating task. In contrast, German speakers show a strong subject-before-object preference (e.g., Bader & Häussler, 2010; Bornkessel et al., 2002), a preference which our materials conformed to. In addition, the case cue in our Experiment 2 was more prominent than in Hopp's study because within the first postverbal argument, case was marked both on the prenominal article and on the adjective following it. This may have increased the likelihood of the L2 speakers processing case cues successfully in our study.

The results from Experiment 2 indicate that highly proficient L2 speakers can indeed use morphosyntactic information predictively under certain conditions. Note, however, that in our experimental materials, the first post-verbal argument either consisted of an entity that could either receive something, thus functioning as a plausible recipient, or of an entity that could be given/presented to someone, thus functioning as a plausible theme. Since the animacy of the first postverbal argument varied across experimental conditions and items, it is possible that semantics also had an influence on participants' gaze patterns. Post-hoc analyses reported in Schlenter (2019) indicate that prediction was easier when thematic roles were not only morphosyntactically marked but if a theme or recipient argument also had prototypical animacy features. Future research might want to investigate the interaction of case marking and animacy in L2 prediction more systematically. In addition, comparing L2 groups from typologically different L1 backgrounds might shed more light on the role of L1-L2 similarity (see also Foucart, this volume).

Our findings also have theoretical implications. While the finding that L2 speakers were less certain than L1 speakers in anticipating an upcoming argument in Experiment 1 could possibly be attributed to (slowed) lexical processing (e.g., Hopp, 2018), other findings cannot. These include the observed difficulty with information integration after encountering a non-canonical argument order. While the L1 group rapidly recovered from initial argument competition and correctly identified the final argument in the non-canonical condition, competition between the accusative object/theme and dative object/recipient persisted in the L2 group. This shows that the L2 group was affected more strongly by our word order manipulation (i.e., by non-canonicity) than was the L1 group, which indicates an over-reliance on word order cues during L2 sentence interpretation. The L2 group's performance in Experiment 2 is reminiscent of what Gerth et al. (2017) observed in a self-paced reading experiment on locally ambiguous ('garden-path') sentences: Even though the L2 speakers Gerth et al. tested showed evidence of noticing disambiguating case information, this did not necessarily lead to greater comprehension accuracy of non-canonically ordered German sentences. L2 speakers' apparent difficulty integrating competing cues that was observed in ours and other studies supports claims to the effect that L1 and L2 speakers differ in their relative weighting of information sources during processing (Clahsen & Felser, 2006, 2018; Cunnings, 2017; compare also Grüter et al., 2020, for evidence from the processing of Mandarin classifiers). Another possible explanation for this finding was pointed out to us by an anonymous reviewer: L2 speakers might have had difficulty recovering from an initially incorrect analysis (e.g., Hopp, 2015; Pozzan & Trueswell, 2016), that is, from assuming that arguments were ordered canonically when in fact they were not. Note that the above possible explanations are not mutually exclusive.

We conclude that under certain (favorable) conditions L2 speakers can predict across different linguistic domains. L2 morphosyntactic prediction ability is not necessarily more limited than semantic prediction ability. Nevertheless, L2 prediction ability might be somewhat more limited in comparison to L1 prediction, with L2 speakers showing greater difficulty with information integration than L1 speakers.

Acknowledgments

We like to thank Cecilia Puebla Antunes for her help with the picture stimuli, and the audiences at the ISBPAC 2018 and the PIPP 2019 conference for helpful comments and stimulating discussion.

Funding

This research was supported by an Alexander-von-Humboldt professorship awarded to Harald Clahsen.

References

Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua*, 120(3), 717–762. https://doi.org/10.1016/j.lingua.2009.05.007

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Barr, D. J., Gann, T. M., & Pierce, R. S. (2011). Anticipatory baseline effects and information integration in visual world studies. *Acta Psychologica*, 137(2), 201–207. https://doi.org/10.1016/j.actpsy.2010.09.011

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot International*, 5(9/10), 341–345.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition*, 85(2), B21–B30. https://doi.org/10.1016/S0010-0277(02)00076-8

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 1029–1040. https://doi.org/10.1037/a0015901

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42. https://doi.org/10.1017/S0142716406060024

Clahsen, H., & Felser, C. (2018). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition*, 40(3), 693–706. https://doi.org/10.1017/S0272263117000250

Cunnings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, 20(4), 659–678. https://doi.org/10.1017/S1366728916000675

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469

Dink, J. W., & Ferguson, B. (2015). *eyetrackingR: An R library for eye-tracking data analysis*. Retrieved from http://www.eyetrackingr.com

Donnelly, S., & Verkuilen, J. (2017). Empirical logit analysis is not logistic regression. *Journal of Memory and Language*, 94, 28–42. https://doi.org/10.1016/j.jml.2016.10.005

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). Multipic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, 71(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915

Frenck-Mestre, C., Kim, S. K., Choo, H., Ghio, A., Herschensohn, J., & Koh, S. (2019). Look and listen! The online processing of Korean case by native and non-native speakers. *Language, Cognition and Neuroscience*, 34(3), 385–404. https://doi.org/10.1080/23273798.2018.1549332

Gerth, S., Otto, C., Felser, C., & Nam, Y. (2017). Strength of garden-path effects in native and non-native speakers' processing of object–subject ambiguities. *International Journal of Bilingualism*, 21(2), 125–144. https://doi.org/10.1177/1367006915604401

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, 35(2), 221–234. https://doi.org/10.1080/23273798.2019.1648840

Häussler, J., & Bader, M. (2011). Grammar- versus frequency-driven syntactic ambiguity resolution: The case of double-object constructions. In M. Lamers & P. de Swart (Eds.), *Case, word order and prominence: Interacting cues in language production and comprehension* (pp. 273–301). Springer.

Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, 17, 5–27. https://doi.org/10.11431/secondlanguage.17.0_5

Hopp, H., & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, 40(1), 171–199. https://doi.org/10.1017/S0272263116000437

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, 98, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, 53(4), 372–380. https://doi.org/10.3758/BF03206780

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Mitrofanova, N., Sekerina, I. A., Sauermann, A., Gagarina, N., & Westergaard, M. (2019, June 27). Cross-linguistic differences in predictive use of grammatical case in Russian and German (comparing monolingual and bilingual children) [Conference presentation]. International Symposium on Bilingualism (ISB), University of Alberta, Canada.

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881

Peters, R. E., Grüter, T., & Borovsky, A. (2015). Anticipatory and locally coherent lexical activation varies as a function of language proficiency. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1865–1870). Cognitive Science Society.

Pozzan, L., & Trueswell, J. (2016). Second language processing and revision of garden-path sentences: A visual world study. *Bilingualism: Language and Cognition*, 19(3), 636–643. https://doi.org/10.1017/S1366728915000838

R Core Team. (2020). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Schlenter, J. (2019). Predictive language processing in late bilinguals: Evidence from visual-world eye-tracking (Doctoral dissertation). University of Potsdam. publish.UP. https://doi.org/10.25932/publishup-43249

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

CHAPTER 4

Influence of syntactic complexity on second language prediction

Eunjin Chun, Si Chen, Shulin Liu and Angel Chan
The Hong Kong Polytechnic University

This study investigated the influence of syntactic complexity on prediction in second language (L2) processing. In a visual world eye-tracking experiment, we compared L2 listeners' prediction while processing simple (e.g., *The dancer will open/ get the present*) vs. complex sentences (e.g., *I know the friend of the dancer that will open/ get the present*). Prediction was measured by comparing fixations to targets (e.g., present) between semantically biasing (e.g., *open*) vs. neutral verb (e.g., *get*) conditions. Results showed that L2 speakers generated predictions while processing complex as well as simple sentences. However, the prediction effect during complex sentence processing emerged somewhat later. These findings suggest that L2 prediction is influenced by syntactic complexity which can increase cognitive load during sentence processing.

Introduction

Research suggests that first language (L1) speakers are likely to predict upcoming linguistic information during comprehension (see Kuperberg & Jaeger, 2016, for a recent review). Findings of linguistic prediction have advanced our understanding of and provided important theoretical implications for language processing and learning (e.g., Chang et al., 2006; Pickering & Garrod, 2013). Specifically, proactive anticipation of upcoming linguistic information suggests that top-down processing is much more engaged in comprehension than it was traditionally considered (Marslen-Wilson, 1973). In addition, as prediction facilitates comprehension, predictive processing could partly account for L1 speakers' rapid comprehension with great ease (Altmann & Kamide, 1999; Altmann & Mirković, 2009; Federmeier, 2007; Kamide et al., 2003). Finally, linguistic prediction has provided supporting evidence for learning accounts which propose that prediction mechanisms underlie implicit learning (e.g., the error-based learning account; Chang et al., 2012). Under these learning accounts, learning is claimed to occur in the process of reducing

https://doi.org/10.1075/bpa.12.04chu

prediction error and thus predictive processing is considered vital for learning (for discussion, see Hopp, this volume).

Given these theoretical implications, it is no surprise that a great deal of attention has been paid to prediction in second language (L2) learners as well. L2 speakers' comprehension has typically been shown to be slower and more difficult than L1 speakers' comprehension, and this could possibly be related to L2 speakers' lack of predictive abilities (Brouwer et al., 2017). Accordingly, research on L2 prediction has primarily focused on L2 speakers' predictive ability, namely whether they are able to predict and if so, whether they can predict to the same extent as L1 speakers (Dijkgraaf et al., 2017; Grüter et al., 2012; Grüter et al., 2014; Martin et al., 2013). Such studies manipulated different types of predictive cues and reported findings that L2 speakers are able to predict (Chambers & Cooke, 2009; Dijkgraaf et al., 2017; Koehne & Crocker, 2015). However, L2 speakers' engagement in prediction seems to vary depending on the types of predictive cues, interacting with multiple factors that influence L2 processing in general (e.g., age of acquisition, L2 proficiency, and cross-linguistic differences between L1 and L2). These interactions remain rather unclear, as do the factors modulating L2 speakers' predictive processing. Exploring the interactions or the mediating factors will provide a better picture of linguistic prediction and ultimately help us better understand predictive mechanisms as well as develop theoretical accounts of linguistic prediction. The current study therefore aims to contribute to this strand of research by investigating whether L2 prediction is influenced by syntactic complexity.

Prediction in comprehension

There has been a surge of research on linguistic prediction during the last two decades, and such studies yielded the consensus that L1 speakers – children as well as adult speakers – make predictions about upcoming linguistic input during comprehension (Altmann & Kamide, 1999; Kuperberg & Jaeger, 2016). Prediction in this paper is defined as pre-activation of upcoming linguistic input (Huettig & Guerra, 2019; Pickering & Gambi, 2018). Suppose that a comprehender processes a sentence like *On his birthday, the boy cut the cake*. If comprehenders predict 'cake', some linguistic information regarding *cake* (e.g., conceptual feature +EATABLE, some phonological feature /keɪk/, and grammatical gender information of *cake* if the language marks gender) is pre-activated upon reading or listening to *cut the* (i.e., before they read or listen to the word *cake*).

Pre-activation of such linguistic information has predominantly been measured using electrophysiological responses or eye movements. Using the phonological regularity in English (e.g., *a* + consonant-initial words and *an* + vowel-initial words), DeLong et al. (2005) designed an eletroencephalography (EEG) experiment to measure pre-activation of specific articles and nouns. Native English participants read sentences of varying constraint (e.g., *The day was breezy so the boy went outside to fly…*) that included target articles and nouns with ranges of cloze probability (e.g., *a kite* is highly likely whereas *an airplane* is less likely in the given example sentence). When participants were expecting *a kite* but *an airplane* was presented, the violation of the phonological regularity would elicit a change in the amplitude of the ERPs for the articles. As expected, the amplitude of the N400 elicited by the articles varied depending on article expectancy. The N400 was larger when the cloze probability of the article and noun was smaller. Since the articles were grammatically and semantically congruent within the contexts, the negative correlation between the N400 amplitude at the article and the article cloze probability was interpreted as suggesting that L1 readers make probabilistic predictions about upcoming nouns and pre-activate some phonological information of the words. That is, as they anticipate the phonological form of a noun (e.g., a consonant-initial word, *kite*) and expect one article (e.g. *a*) relative to the other (e.g., *an*), they would experience integration difficulty when the less-expected article was presented (see Nieuwland et al., 2018, for different findings).

Stronger evidence for prediction comes from visual world eye-tracking studies which measure listeners' anticipatory looks to upcoming referents in a visual display. For example, Altmann and Kamide (1999) recorded eye movements from L1 English speakers while they were listening to sentences containing semantically restrictive verbs (e.g., *The boy will eat the cake*) or neutral verbs (e.g., *The boy will move the cake*). When presented with a scene depicting a cake and inedible objects, these English speakers' fixations on the cake increased as soon as the constraining verb was heard. Crucially, participants showed more fixations on the cake than the other objects even before they heard the direct object *cake*. These anticipatory eye movements suggest that L1 listeners predict plausible direct objects using the selectional restrictions of verbs. To date, visual world eye tracking studies have shown that L1 speakers use various types of cues to predict upcoming information (e.g., semantic cues: Altmann & Kamide, 1999; Kamide et al., 2003; syntactic cues: Arai & Keller, 2013; discourse cues: Otten et al., 2007; prosodic cues: Ito & Speer, 2008; Nakamura et al., 2012; Perdomo & Kaan, 2021).

Prediction in L2 comprehension

Prediction is not only resource-expensive processing, but it also entails the risk of failure. Building up expectations during computations of ongoing information already requires more cognitive resources than simply integrating information which unfolds in a rapid fashion. It demands further cost if the predictions do not match the actual linguistic input and thus need to be rapidly revised while the processor also needs to keep up with the ongoing information. Considering the high expense, generating predictions can be challenging or risky for L2 speakers and may not be possible for those who are already overburdened by online integrative processing during comprehension. Therefore, L2 speakers have been claimed to have Reduced Ability to Generate Expectations (the RAGE Hypothesis; Grüter et al., 2014). Studies on L2 prediction revealed that the influence of RAGE on L2 processing differs depending on linguistic cues, interacting with various factors that influence L2 processing in general.

When semantic cues are available in the context, L2 speakers are likely to use them predictively, often to a similar extent as L1 speakers. Chambers and Cooke (2009) replicated Altmann and Kamide's (1999) semantic prediction effect with L2 speakers. When presented with French sentences like *Marie va nourrir la poule* ('Marie will feed the chicken') or *Marie va décrire la poule* ('Marie will describe the chicken'), late L2 learners of French with high proficiency showed more anticipatory looks to the target picture of *poule* ('chicken') upon hearing the verb *nourrir* ('feed') compared to hearing the verb *décrire* ('describe'). Such semantic cues could be readily used for prediction by L2 speakers even at the beginning level (Koehne & Crocker, 2015). Unbalanced bilinguals also showed semantic prediction in both L1 and L2, and their semantic prediction in L1 did not differ from monolinguals' prediction (Dijkgraaf et al., 2017). Importantly, semantic prediction in L2 has been shown to be affected by the spread of semantic activation (Chambers & Cooke, 2009; Dijkgraaf et al., 2019). In a study using a visual world paradigm, Dijkgraaf et al. (2019) tested whether bilinguals make semantic predictions to the same extent in their L1 and L2. When presented with a display containing a picture of either a target or a semantic competitor and three unrelated objects, bilinguals showed more anticipatory fixations on the target and the semantic competitor than on the other objects in both L1 and L2. They also showed more anticipatory fixations to semantically related competitors than the other objects, and this effect appeared stronger and earlier in their L1 than in L2. These findings indicate that semantic predictions in L2 are made in a similar fashion as those in L1, but the extent of semantic prediction seems to be influenced by the spread of semantic activation. L2 speakers' inevitable experience of lexical competition through cross-lingual word coactivation and their lower language experience in L2 could result in weaker semantic representations, which in turn may be attributable to weaker and slower semantic activation during L2 semantic prediction.

As for (morpho)syntactic cues, prior studies revealed mixed findings with much more variance. Some studies found that L2 speakers show difficulties in using (morpho)syntactic information for prediction (e.g., Grüter et al., 2012;

Lew-Williams & Fernald, 2010; Martin et al., 2013; Mitsugi & MacWhinney, 2016) whereas others demonstrated that L2 speakers could use this type of information predictively (e.g., Dussias et al., 2013; Foucart et al., 2014; Schlenter & Felser, this volume). The picture emerging from recent studies is that syntactic prediction in L2 interacts with factors such as L2 speakers' proficiency and their L1 backgrounds. For instance, intermediate L2 Japanese learners who had grammatical knowledge about Japanese case markers could not utilize the case marking information to predict an upcoming word (Mitsugi & MacWhinney, 2016). Similarly, moderately proficient English learners of Spanish could not use grammatical gender information for prediction despite their relevant grammatical knowledge. In contrast, highly proficient L2 speakers could use grammatical gender information for prediction in a native-like way (Dussias et al., 2013). L1 backgrounds also seem to have a significant influence on syntactic prediction in L2. In contexts where the L1 shared similar syntactic properties with the L2, L2 speakers could use syntactic information to make predictions. van Bergen and Flecken (2017) tested three groups of Dutch learners with different L1 backgrounds (English, French and German), and only those with German L1 background, whose L1 similarly encodes object position, could make use of Dutch placement verbs to predict object position. This pattern was observed regardless of age of acquisition. Both early and late L2 learners could make syntactic predictions in a similar manner to L1 speakers when the same syntactic feature (e.g., gender agreement rule) exists in L1 and L2 (Foucart et al., 2014; see also Foucart, this volume).

Variation in prediction and mediating factors

As introduced above, L2 speakers show large variability in prediction. However, the case is not limited to L2 speakers. The current literature reveals variation in prediction within and across populations (Federmeier, 2007; Federmeier et al., 2002; Huettig & Guerra, 2019). Considering that prediction is basically part of language processing, which is constrained by timing, whatever influences language processing in general can be considered to give rise to variation in prediction within and across speakers. Exploring this variation will help us further understand how prediction mechanisms work.

Studies showing when predictions fail provided as much information about predictive mechanisms as ones showing when predictions succeed. For instance, Chow et al. (2018) investigated the impact of argument role information on verb predictions using the *ba* construction in Mandarin Chinese. The particle *ba* is always positioned between the subject and the direct object (i.e., subject + *ba* + object + verb). Therefore, syntactic roles of preverbal arguments are easily noticeable in this construction, and the likelihood of an upcoming verb is changed by the arguments' structural roles. That is, the verb, *arrest* is more likely to appear in a sentence like *Jingcha ba xiaotou …* ('cop BA thief …') than in a sentence like *Xiaotou ba jingcha …* ('thief BA cop …'). In an ERP study using this *ba* construction, Chow et al. manipulated argument role information (canonical vs role-reversed sentences) and predictability (high vs. low predictability) in Experiment 1, and the linear distance between the pre-verbal arguments and the verb (e.g., '*cop ba thief (yesterday after noon) arrest*') in Experiment 2. Results of this study showed that L1 comprehenders failed to predict verbs using argument role information (i.e., no N400 effect by argument role reversals even in the high predictability condition), but their verb predictions were sensitive to this information when the context was highly predictable and more time was available (i.e., an N400 effect by argument role reversals in the the long-distance condition with high predictability). The delayed impact of argument role information on verb prediction indicates that verb predictions evolve over time. Chow et al. interpreted these findings as suggesting that prediction involves computations requiring different amounts of time. In addition, Huettig and Guerra (2019) observed that L1 prediction is constrained by experimental conditions such as speech rate, preview time of visual context, and participant instructions. While listening to normal speech, L1 comprehenders showed prediction effects only when they were given extensive preview time (4 sec), but not when they had short preview time (1 sec). Also, they showed only a small prediction effect under the condition with a normal speech rate and a short preview even though they were explicitly instructed to predict. These results provided evidence against the notion that human brains are essentially prediction machines (Clark, 2013).

Furthermore, variation in prediction prompted research exploring potential factors that influence predictive processing and helped us understand which factors play an important role in prediction. According to Huettig and Janse (2016), L1 adult speakers' syntactic prediction is modulated by their working memory abilities and processing speed. The better working memory and the faster processing speed, the more likely that L1 comprehenders predict (but see Otten & Van Berkum, 2009, for a failure of finding effects of working memory span on prediction). Regarding inconsistent prediction effects in L2, Mitsugi and MacWhinney (2016) suggested that this may be due primarily to limited cognitive resources in L2 speakers. During sentence comprehension, speakers need to retrieve relevant information from memory and integrate this information as the sentence unfolds. They also need to integrate non-linguistic information from the language environment with the linguistic information that they retrieved from memory. In this process, if L2 speakers use up cognitive resources for complex online computations, few resources may be left for prediction (see Ito and Pickering, this volume). Given that greater cognitive resources are required for L2 comprehension than L1 comprehension (Segalowitz

& Hulstijn, 2005), L2 speakers may have difficulties in accessing necessary knowledge for prediction (e.g., grammatical knowledge) particularly when they are given complex cues (e.g., combinations of semantic and syntactic information).

This view is in line with the RAGE hypothesis under which the limitation of L2 prediction is because of cognitive burden that L2 speakers experience during online computation of linguistic input. This issue was more directly addressed in some recent studies. With the hypothesis that increased cognitive load would interfere with prediction, Ito et al. (2018) investigated the relationship between cognitive load and predictive eye movements in both L1 and L2 speakers. In visual world eye-tracking experiments, half of the participants in each speaker group listened to sentences with a simple SVO structure (similar to those used in Altmann & Kamide, 1999) and clicked a mentioned object on the visual displays (i.e., listen-and-click task only). The other half performed an additional working memory task. Compatible with previous findings, both groups of participants who did the listen-and-click task only showed more anticipatory fixations on the plausible objects (e.g., scarf) when listening to semantically restrictive verbs (e.g., *fold*) than when listening to semantically neutral verbs (e.g., *find*). However, this semantic prediction effect was delayed for those who performed the concurrent working memory task in both L1 and L2 speaker groups. These findings were taken to support the view that an additional working memory task can impose a cognitive load even during simple sentence processing and consequently delay the prediction effect.

However, the study by Ito et al. lacked ecological validity in that the cognitive load externally imposed by the memory task is far from the cognitive challenges that speakers usually experience during sentence processing. The effects of cognitive load on prediction can be investigated in a more natural setting by manipulating syntactic complexity. L2 speakers have shown difficulties when they process complex sentences, and their processing is modulated by cognitive capacities (Dussias & Piñar, 2010; Zhou et al., 2017). These findings imply that syntactic complexity is one of the factors which can increase cognitive load during sentence processing. In this regard, Chun and Kaan (2019) investigated whether L2 speakers are able to make predictions even while processing complex sentences. Thus far, semantic prediction during simple sentence processing has been well-attested, and therefore their study solely focused on the comparison of prediction during complex sentence processing between L1 and L2 speakers. In a visual world eye-tracking experiment, participants listened to relative clause (RC) sentences with complex noun phrases containing either semantically biasing or neutral verbs (e.g., *I know the friend of the dancer that will open/ get the present*). As processing complex sentences would increase cognitive load and more cognitive resources are required for L2 processing than L1 processing, L2 speakers were expected to feel overburdened by online computation itself and fail to make predictions during complex sentence processing.

Contrary to this expectation, L2 speakers showed native-like prediction, directing their eyes to the predictable target object (e.g., present) based on the semantic information of the verbs while processing complex sentences.

Despite the novel finding of L2 prediction effects during complex sentence processing, Chun and Kaan's (2019) study had some limitations. They did not include comprehension questions for the sake of examining prediction during natural sentence processing. As participants' comprehension was not probed, the authors could not exclude the possibility that L2 participants did not fully parse the structures (i.e., shallow processing). If this was the case, those who had not fully parsed the structures might have not used up resources for online computation, which could have led them to use resources for prediction. In other words, L2 participants could just focus on the verb information (e.g., *open*) and anticipate the plausible direct objects (e.g., *present*), without attaching the RC to any of the noun phrases (NPs). They then might not have experienced much cognitive load and used resources for prediction even during complex sentence processing. In addition, Chun and Kaan compared prediction effects between L1 and L2 speakers only during complex sentence processing, and their results did not inform us of the extent to which syntactic complexity can influence L2 prediction. The present study was therefore designed to extend Chun and Kaan's (2019) study by ruling out the possibility of incomplete parsing (using comprehension questions) and including a simple sentence condition.

The current study

In this study, we conducted a visual world eye-tracking experiment to investigate the influence of syntactic complexity on L2 prediction. We included comprehension questions to encourage L2 speakers' full parsing. We also compared L2 prediction during simple vs. complex sentence processing to understand the extent to which syntactic complexity influences L2 prediction.

To replicate L2 prediction during complex sentence processing, we used the same sentences as those in Chun and Kaan (2019) for the complex sentence condition. We manipulated semantic associations between the critical elements (e.g., between the agent and the verb, and between the verb and the theme) to provide semantic cues. We created the materials for the simple sentence condition (e.g., *The dancer will open the present*) by extracting critical elements from the complex sentences (e.g., *I know the friend of the dancer that will open the present*). This was to keep the semantic cues consistent between the two sentence conditions. In this way, the materials in both sentence conditions could only differ in terms of syntactic complexity while keeping all the key lexical items identical. However, the same lexical items between the two sentence conditions could yield repetition effects, and thus we employed a between-subject design. Based on the previous findings that cognitive load delays prediction effects (Ito et al., 2018), we expected L2 prediction to be delayed during complex sentence processing if syntactic complexity increases cognitive load for online computations and in turn influences prediction.

Method

Participants

Fifty Chinese learners of English were recruited from a university in Hong Kong. They participated in this study for monetary compensation ($80 HKD per hour). Half of the participants were exposed to the simple sentences (simple sentence group) and the other half to the complex sentences (complex sentence group). None of these participants had hearing problems or learning disorders. Before the main experiment, all participants completed informed consent forms approved by the Institutional Review Board of the university. A battery of tests was used to examine participants' working memory capacity and linguistic proficiency: a reading span task (Kane et al., 2004), the Shipley vocabulary test (Shipley, 1940), and the grammar and cloze section of the MELICET (Michigan English Language Institute College English Test; Michigan English Language Institute, 2001). We also collected self-reported scores from the IELTS (International English Language Testing System) as an additional proficiency measure. Participant information is provided in Table 1. The two groups of L2 participants did not differ in working memory capacity (the reading span task: $t(47.23) = -0.12$, $p = .91$), vocabulary size (the Shipley test; $t(47.62) = 1.76$, $p = .08$) or any measures of proficiency (IELTS: $t(41.69) = -0.67$, $p = .51$; the Grammar and Cloze section of MELICET: $t(48) = 0.46$, $p = .65$).

Table 1. Participant information

	Simple sentence group	**Complex sentence group**
Age	*M*: 24.24 (*SD*: 2.09)	*M*: 24.48 (*SD*: 2.35)
IELTS	*M*: 6.56 (*SD*: 0.17)	*M*: 6.60 (*SD*: 0.25)
Shipley	*M*: 21.32 (*SD*: 3.68)	*M*: 19.56 (*SD*: 3.37)
MELICET	*M*: 30.32 (*SD*: 6.14)	*M*: 29.52 (*SD*: 6.17)
Reading Span	*M*: 29.00 (*SD*: 7.75)	*M*: 29.24 (*SD*: 6.81)

Materials

In order to assess participants' use of semantic cues for prediction, half of the sentences in each sentence condition included semantically biasing verbs and the other half included neutral verbs. The experimental stimuli for the complex sentence condition consisted of fourteen pairs of sentences including object-modifying and subject-extracted relative clauses (e.g., *I know the friend of the dancer that will open/ get the present*). As for the experimental stimuli for the simple sentence condition, fourteen pairs were created by extracting the second noun phrase (NP2), the verb, and the object in the RCs from the complex sentences (e.g., *The dancer will open/get the present*). That is, the simple sentences contained exactly the same verbs and objects as those in the RCs from the complex sentences, and the subjects of the simple sentences always corresponded to the NP2 of the complex sentences. In this way, the semantic information for predictive cues was controlled for between the two sentence conditions (simple vs. complex). Sixteen RC sentences with one NP (e.g., *The chef knows the girl that will cook the chicken*) were prepared for fillers and used for both conditions.

The complex sentences were recorded by a female native American English speaker at the sampling rate of 44.1 kHZ. They were recorded at a rate of 3.1 words per second. The duration of the verb and the determiner was kept constant across the items (i.e., verb + *the*: 642 ms). This was to provide participants with the same amount of time to make predictive eye movements. Then, the auditory stimuli for the simple sentence condition were spliced from those for the complex sentence condition. In this way, not only the speech style and rate but the duration of the critical time window for prediction was also consistent across the items in both conditions. For the visual displays, we prepared fourteen scenes that depicted three objects (e.g., a target and two distractors) and two agents (e.g., a boy and a girl) referring to the two noun phrases in the complex sentences (see Figure 1). The item which can be an appropriate theme of the semantically biasing verb was coded as the target. The two distractors were as likely to be the themes of the semantically neutral verb as the target. For example, the target *present* in Figure 1 is the only item that can be the theme of *open* whereas the distractors *money* and *trophy* can be as likely to be the themes of *get* as the target *present*. To verify this semantic manipulation, a norming study was separately conducted with native speakers of American English (N = 101). Participants were asked to complete a given sentence fragment (e.g., *The boy will open/ get* ______) using any object on the visual display. In this test, participants chose the target item 96% of the time when the biasing verb was presented whereas they did so 42% of the time when the neutral verb was presented. The locations of the targets and the distractors were randomized on each trial.

Figure 1. An example visual display

We prepared comprehension questions (e.g., *Who will open the present?*) to encourage participants to fully parse the experimental sentences. Since there were no correct answers for the ambiguous RCs in the complex sentence condition, participants' parsing accuracy could not be checked. However, research has shown that speakers have attachment bias when parsing ambiguous RC sentences. Thus, we prepared a quick behavioural task to identify each participant's attachment bias and compared their attachment interpretations in the behavioral task with their comprehension answers during the eye-tracking task. For this behavioral task, we prepared four sets of auditory stimuli with each set consisting of nine ambiguous RC sentences (e.g., *Michelle sees the child of the mother that is talking to the woman*) and twelve fillers (RC sentences with one NP). Participants' comprehension answers during the eye-tracking task were expected to be parallel to their RC attachment interpretations in the behavioral task on the assumption that they would use their attachment bias when processing the ambiguous RCs in the complex sentence condition.

Procedure

The behavioral task to identify participants' attachment bias was first administered using E-prime (Psychological Software Tools). The lists of ambiguous RC sentences were counterbalanced across participants and the experimental sentences in each list were pseudorandomized with the fillers intervening one or two experimental sentences. In this behavioral task, participants were instructed to listen to auditory sentences (e.g., *Michelle sees the child of the mother that is talking to the woman*) via headphones and answer questions (e.g., *Who is talking to the woman?*) by pressing a button which corresponded to the first noun phrase (NP1) or the second noun phrase (NP2).

Then, the visual world eye-tracking task was conducted on an Eyelink 1000 system with a chin rest. Before the beginning of the experiment, we completed an automatic 9-point calibration and validation routine using a standard black and white bull's eye image. Recalibration was conducted during the experiment whenever necessary. The visual displays were presented using a PC computer running EyeLink Experiment Builder software (SR Research, Mississauga, Ontario, Canada) and auditory stimuli were presented using the same computer via head phones. While listening to auditory sentences, participants' eye movements were recorded at a 500 Hz sampling rate. Before the main experiment, participants practiced with 5 trials and the practice was repeated until they understood the task.

Each trial started with a bull's eye image at the center of the screen which served as a drift correction. Participants were instructed to fixate on it and press the space bar whenever they were ready to proceed. Once the space bar was pressed, a visual display was presented for 2000 ms before the onset of a sentence. After this preview, participants heard auditory sentences and clicked the last-mentioned object from the auditory stimuli. The visual display remained on the screen until they clicked. The behavioral mouse-clicking was to encourage participants to pay attention to the task and look at the visual display. Finally, participants answered comprehension questions by pressing buttons. The same questions were presented for both sentence conditions (e.g., *Who will open the present?*), but the questions were used to probe RC attachment interpretations for the complex sentence condition. The trials of the eye-tracking task were randomly presented.

Results

Behavioral task accuracy

The accuracy of the mouse-clicking responses was 99.3 % in the simple condition and 93.4 % in the complex condition. With regards to comprehension accuracy, participants in the simple sentence condition showed 96.7% accuracy. For the complex sentence condition, participants' comprehension answers were significantly correlated with their attachment interpretations in the behavioural task ($r(23) = .43$, $p = .03$). This result suggests that participants in the complex sentence condition parsed the ambiguous RC structures following their attachment bias.

Eye-tracking data analysis

The eye-tracking data were analyzed using the lme4 package (Bates et al., 2015) in R (R Core Team, 2016). First, we preprocessed the eye-tracking data using the VWPre package (Porretta et al., 2017). The fixation proportions on the target and the other objects were calculated for each 20 ms time bin relative to the onset of the target noun, and then fixation probability in every time bin was transformed into log odds, using the empirical logit function (Barr, 2008) installed in the package. Track loss or blinks were not included for fixations, and the trials with incorrect mouse-clicking responses and comprehension answers were excluded for the eye-tracking analysis.

Figure 2 plots the fixation proportions to the targets between the biasing (the solid lines) and the neutral verb conditions (the dotted lines) in the simple and complex sentence conditions. This time-course plot shows mean fixation proportions to the targets from 1000 ms before to 1000 ms after the onset of the target noun. The plot was time-locked to the onset of the target noun (the vertical solid line at time zero) and the vertical dashed line marks the onset of the verb in the spoken sentence. The error-bands indicate 95% bootstrapped confidence intervals of mean fixation proportions to the targets. As seen in Figure 2, L2 participants in both simple and complex sentence conditions showed more anticipatory looks to the targets in the biasing versus the neutral verb conditions before they heard the target noun.

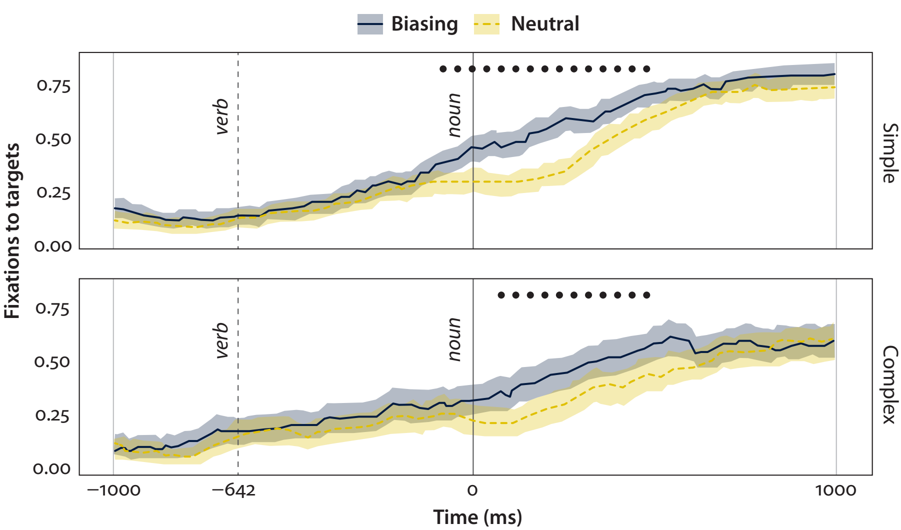

Figure 2. Mean fixation proportions to the target objects in the two verb conditions (biasing vs. neutral) in the simple sentence condition (top panel) and in the complex sentence condition (bottom panel)

Note. Error-band: 95% bootstrapped confidence intervals. A solid circle (•): significant differences in fixations on the targets between the two verb conditions in each time bin (|t|s > 2).

For statistical analysis of the prediction effect, we constructed a linear mixed effects model over the log-transformed fixation probabilities on targets from 442 ms before to 200 ms after the target noun onset. To meet the statistical assumption that the dependent variable has an unbounded range, we used the log-transformed fixation probabilities as the dependent variable (Barr, 2008). The analysis window was set considering a latency of 200 ms for eye movement planning (Matin et al., 1993), and the duration of the verb and the determiner (642 ms). The eye movements during this time window would capture prediction using semantic information of the verb before encountering the noun. For the fixed effects, the model included contrast coded verb type (neutral verb coded as −0.5 vs. biasing verb coded as 0.5), sentence type (simple sentence coded as −0.5 vs. complex sentence coded as 0.5), and an interaction between verb type and sentence type. For the random effects, the model included random intercepts for participants and items, and verb type as a by-participant random slope. The final model did not include a by-item random slope for verb type because the model with it did not converge (see the summarized results of the final model in Table 2). In support of the fixation differences between the two verb conditions as shown in Figure 2, L2 participants' anticipatory fixations onto the targets were significantly influenced by verb type ($b = 0.55$, $SE = 0.14$, $t = 3.96$, $p < .001$). Participants showed more anticipatory fixations onto the targets as soon as they heard the biasing verb rather than the neutral verb (mean fixation proportions to the targets: 0.32 for the simple sentence condition and 0.28 for the complex sentence condition).

Table 2. Summary of the mixed effects model as a function of verb type (Verb) and sentence type (Sentence)

	b	SE	t	p
Intercept	−1.38	0.15	−8.91	< 0.001
Sentence	−0.08	0.20	−0.38	0.70
Verb	0.55	0.14	3.96	< 0.001
Sentence: Verb	−0.15	0.20	−0.74	0.46

However, there was neither a main effect of sentence type nor an interaction between sentence type and verb type. This may be because the time window collapsed over for the analysis was so wide that potential differences between the two sentence conditions may have been obscured. We therefore conducted a time-course analysis following previous work (Borovsky et al., 2012; Ellis et al., 2015; Ito et al., 2018). A separate model for each sentence condition was run for every 40 ms bin from 442 ms before to 500 ms after the target noun onset. The model included the fixed factor of verb type, random intercepts for participants and items, and verb type as a by-participant random slope. Since this type of time-course analysis can increase the likelihood of Type I errors, researchers have been conservative when reporting results (Baayen, Davidson, & Bates, 2008). They typically report results showing consistent reliability with the absolute t-value exceeding 2 over multiple bins (e.g., more than three to five consecutive bins). Significance of the time-course analysis is shown on the top of the graphs in Figure 2. A solid circle (•) indicates significant differences in fixations on the targets between the two verb conditions in each time bin ($|t|$s > 2). Results of this analysis showed that prediction effects consistently appeared from 82 ms before the target noun onset onwards in the simple sentence condition, but they appeared from 78 ms post target noun onset onwards in the complex sentence condition.

Discussion

This study investigated the influence of syntactic complexity on predictive eye movements in L2 speakers. In a visual world eye-tracking experiment, half of the participants listened to simple sentences and the other half listened to complex sentences. For each sentence condition, half of the sentences contained semantically biasing verbs and the other half contained semantically neutral verbs. The results showed that L2 listeners made more anticipatory looks to the target object (e.g., present) upon hearing the biasing verbs (e.g., *open*) than the neutral verbs (e.g., *get*) regardless of sentence type. However, participants who listened to the complex sentences showed prediction effects somewhat later than those who listened to the simple sentences. Taken these findings together, L2 speakers are able to make use of semantic information to predict during complex sentence processing as well as simple sentence processing, but syntactic complexity may delay L2 prediction.

The semantic prediction effect during simple sentence processing is consistent with previous findings. Similar to L2 participants in Chambers & Cooke (2009), and Dijkgraaf et al. (2017), L2 speakers in this study showed semantic prediction effects. It has been consistently reported that L2 speakers generate predictions while processing syntactically simple sentences (e.g., SVO sentences) when semantic information of the verb is manipulated. As proposed by Ito et al. (2018), this may be because such semantic information is a type of predictive cue that can be used by L2 speakers with relative ease. Computing syntactically simple sentences is not that resource-expensive, and therefore L2 speakers seem to be able to allocate resources for predictive processing. In short, L2 speakers may have enough resources for prediction during simple sentence processing particularly when semantic cues, which can be relatively easy to use, are provided.

In addition, L2 participants who processed the complex sentences in this study showed anticipatory looks to the target object before the noun could have been processed. L2 speakers' predictive abilities during complex sentence processing were reaffirmed by this study in which L2 speakers were encouraged to fully parse sentences by the comprehension questions. The L2 prediction effect during complex sentence processing is therefore not likely to be due to incomplete or local parsing. The significant correlation between participants' comprehension answers and attachment interpretations in the behavioral task supports the idea that participants parsed the ambiguous RCs using their attachment bias. This finding is rather surprising given the processing difficulty that Chinese learners of English would experience when comprehending the complex sentences used in this study. Object-modifying RCs, used in the complex sentence condition, have been shown to be processed more slowly than subject-modifying RCs regardless of extraction type (Gibson et al., 2005) and L1 word order has been found to negatively influence L2 sentence processing when the word order is different between the two languages (i.e., a negative L1 transfer effect on L2 word order processing, Erdocia & Laka, 2018). The participants' L1, Chinese, has the opposite word order (e.g., RC NP1 NP2) from their L2, English (e.g., NP1 NP2 RC), and thus the Chinese learners of English who participated in this study may have experienced more processing difficulties than other L2 speakers whose L1 has the same word order as English.

L2 processing itself requires many resources, so predictive processing in L2 could be hindered by the cognitive load imposed by syntactic complexity if resources were depleted by ongoing complex linguistic computations (Mitsugi & MacWhinney, 2016). However, our findings suggest that proficient L2 speakers may have enough resources to make predictions even while they process syntactically complex sentences, at least the ones with object-modifying and subject-extracted RCs used in this study and in Chun and Kaan (2019). L2 participants in Chun and Kaan's (2019) study were Chinese learners of English who were immersed in an English-speaking country and L2 speakers in this study were Chinese learners of English in Hong Kong who use English as their official language on a daily basis. L2 speakers in these two studies not only showed similar levels of English proficiency, but they were both immersed in English-speaking environments. That is, they may have enough linguistic knowledge and cognitive resources to make predictions during online computations of complex sentences.

Though semantic prediction effects in this study were observed regardless of sentence type, the time-course analysis revealed evidence for prediction during complex sentence processing somewhat later than during simple sentence processing. This finding suggests that cognitive load increased by syntactic complexity may delay predictive processing. As more resources are needed for ongoing computations of complex sentences, fewer resources could be available for predictive processing. This result is therefore compatible with Ito et al.'s (2018) findings that cognitive load delays prediction.

It should also be noted that the results of this study need to be interpreted considering the experimental manipulation. Compared to the visual displays in prior L2 research, the displays in this study were relatively complex. Previously, the visual scenes depicted four objects, one in each quadrant whereas those in this study depicted two agents and three potential themes in a semi realistic background. This visual context can be considered rather complex as a single verb has been found to activate its typical agents and patients (Kukona et al., 2011), and two agents in the visual display were equally possible for the agent of the verb (the main verb in the simple sentences and the verb embedded in the RCs). Taking into consideration its complexity, the visual display in this study was presented for two seconds (Huettig et al., 2011). In addition, participants listened to slow to normal rate of speech. These experimental settings may have enabled participants to fully perceive the visual scenes and process linguistic input to the extent that they could make predictive eye movements towards upcoming objects by integrating linguistic information with the visual scenes. Given recent findings that the speech rate and the preview time for visual displays can influence prediction in L1 comprehension (Huettig & Guerra, 2019), it is possible that predictive eye movements are less likely when L2 comprehenders are given less preview time and/or listen to faster speech. As these are other factors in the linguistic environment which can increase cognitive load during L2 comprehension, it is worth investigating the influence of these factors on L2 prediction.

Finally, this study used the same materials as those in Chun and Kaan (2019) for the complex sentence condition to compare the results between the two studies. Chun and Kaan tried to increase cognitive load during complex sentence processing using sentences containing object-modifying RCs with complex noun phrases. In fact, these sentences are not only syntactically complex but they are also globally ambiguous. The ambiguity driven from the syntactic structure may additionally increase cognitive load. Therefore, participants in the complex sentence condition may have dealt with syntactic complexity coupled with ambiguity. Their delayed prediction effect could be attributable to these factors possibly working together.

In conclusion, we reported findings from a visual world eye-tracking experiment that investigated the influence of syntactic complexity on L2 prediction. We particularly focused on syntactic complexity because it seems to be one of the most commonly encountered challenges in L2 comprehension. L2 speakers often show processing difficulty for syntactically complex sentences. Thus, we expected syntactic complexity to increase cognitive load during sentence processing and in turn influence prediction. The results of this study showed that L2 speakers made predictions on the basis of semantic information while processing complex sentences as well as simple sentences. However, their predictions were somewhat delayed during complex sentence processing. In other words, L2 speakers were able to use semantic cues to predict even under increased cognitive load imposed by syntactically more complex sentences. Yet, prediction was affected by syntactic complexity as we observed some delay in prediction during complex sentence processing. These findings suggest that prediction is resource-constrained and thus L2 speakers' engagement in prediction is mediated by factors such as syntactic complexity that can influence cognitive load during sentence processing. So, it is possible that L2 speakers do not make predictions when comprehending much more syntactically complex sentences than those used in this study (e.g., ones containing embedded object-modifying and object-extracted RCs; *The fact that the president ignored the reporter who the senator attacked on Tuesday bothered the editor*, from Gibson et al., 2005). We call for future studies that investigate how prediction can be modulated by syntactic complexity of different complex constructions. This line of research would not only provide insights into predictive mechanisms but also help us understand why L2 speakers experience more difficulties in learning some syntactic structures than others if prediction mechanisms are indeed related to language learning.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes*, 28(4), 525–560. https://doi.org/10.1080/01690965.2012.658072

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, 112(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Brouwer, S., Özkan, D., & Küntay, A. C. (2017). Semantic prediction in monolingual and bilingual children. In E. Blom, L. Cornips, & J. Schaeffer (Eds.). *Cross-linguistic influence in bilingualism* (pp. 49–74). John Benjamins. https://doi.org/10.1075/sibil.52.04bro

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(4). 1029–1040. https://doi.org/10.1037/a0015901

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Linguistics and Language Compass*, 6(5), 259–278. https://doi.org/10.1002/lnc3.337

Chow, W. Y., Lau, E., Wang, S., & Phillips, C. (2018). Wait a second! delayed impact of argument roles on on-line verb prediction. *Language, Cognition and Neuroscience*, 33(7), 803–828. https://doi.org/10.1080/23273798.2018.1427878

Chun, E., & Kaan, E. (2019). L2 prediction during complex sentence processing. *Journal of Cultural Cognitive Science*, 3(2), 203–216. https://doi.org/10.1007/s41809-019-00038-0

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English second language speakers. *Second Language Research*, 26(4), 443–472. https://doi.org/10.1177/0267658310373326

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915

Ellis, E. M., Borovsky, A., Elman, J. L., & Evans, J. L. (2015). Novel word learning: An eye-tracking study. Are 18-month-old late talkers really different from their typical peers? *Journal of Communication Disorders*, 58, 143–157. https://doi.org/10.1016/j.jcomdis.2015.06.011

Erdocia, K., & Laka, I. (2018). Negative transfer effects on L2 word order processing. *Frontiers in Psychology*, 9, 337. https://doi.org/10.3389/fpsyg.2018.00337

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., McLennan, D. B., de Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39(2), 133–146. https://doi.org/10.1111/1469-8986.3920133

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, 40(5), 1461–1469. https://doi.org/10.1037/a0036756

Gibson, E., Desmet, T., Grodner, D., Watson, D. & Ko, K. (2005). Reading relative clauses in English. *Cognitive Linguistics*, 16, 313–354. https://doi.org/10.1515/cogl.2005.16.2.313

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191–215. https://doi.org/10.1177/0267658312437990

Grüter, T., Rohde, H., & Schafer, A. J. (2014). The role of discourse-level expectations in non-native speakers' referential choices. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the Annual Boston University Conference on Language Development*. (Vol. 1, pp. 179–191). Cascadilla Press.

Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research*, 1706, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137(2), 151–171. https://doi.org/10.1016/j.actpsy.2010.11.003

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58(2), 541–573. https://doi.org/10.1016/j.jml.2007.06.013

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kane, M. J., Tuholski, S. W., Hambrick, D. Z., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2). https://doi.org/10.1037/0096-3445.133.2.189

Koehne, J., & Crocker, M. W. (2015). The interplay of cross-situational word learning and sentence-level constraints. *Cognitive Science*, 39(5), 849–889. https://doi.org/10.1111/cogs.12178

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, 119(1), 23–42. https://doi.org/10.1016/j.cognition.2010.12.002

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244(5417), 522–523. https://doi.org/10.1038/244522a0

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, 53(4), 372–380. https://doi.org/10.3758/BF03206780

Michigan English Language Institute. (2001). *Michigan English Language Institute College English Test*. Ann Arbor, MI: Michigan English Language Institute

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881

Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition*, 125(2), 317–323. https://doi.org/10.1016/j.cognition.2012.07.016

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, 7, e33468. https://doi.org/10.7554/eLife.33468

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. A. (2007). Great expectations: Specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, 8(1), 89–98. https://doi.org/10.1186/1471-2202-8-89

Otten, M., & Van Berkum, J. J. A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, 1291, 92–101. https://doi.org/10.1016/j.brainres.2009.07.042

Perdomo, M., & Kaan, E. (2021). Prosodic cues in second-language speech processing: A visual world eye-tracking study. *Second Language Research* 37(2), 349–375. https://doi.org/10.1177/0267658319879196

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Porretta, V., Kyröläinen, A., van Rij, J., & Järvikivi, J. (2017). VWPre: tools for preprocessing visual world data. R package version, 1.2.3

R Core Team. (2016). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. Retrieved on 15 March, 2020 from: https://www.r-project.org

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). Oxford University Press.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology: Interdisciplinary and Applied*, 9(2), 371–377. https://doi.org/10.1080/00223980.1940.9917704

University of Michigan. (2001). *Michigan English Language Institute College English Test: Grammar, cloze, vocabulary, reading*. University of Michigan.

van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Zhou, H., Rossi, S., & Chen, B. (2017). Effects of working memory capacity and tasks in processing L2 complex sentence: Evidence from Chinese-english bilinguals. *Frontiers in Psychology*, 8, 595. https://doi.org/10.3389/fpsyg.2017.00595

CHAPTER 5

Language prediction in second language

Does language similarity matter?

Alice Foucart
Nebrija University

It is currently accepted that native speakers regularly predict upcoming elements during language comprehension. Empirical evidence suggests that the ability to predict is not as stable in a second language (L2) compared to the first language (L1). One factor that may affect prediction in an L2 is cross-linguistic influence (CLI). Only a handful of studies have tested this hypothesis. Hence, the aim of this chapter is to raise awareness of the importance of CLI, rather than to draw strong conclusions, and to propose directions for future research. After a brief review of CLI on language processing in general, studies that have investigated the circumstances under which these effects may play a (positive/negative) role with regard to prediction are reviewed.

Introduction

It is commonly accepted that native speakers regularly predict upcoming elements during language comprehension. In line with models in cognitive science, we assume that predictions are generated by top-down information and are constantly updated on the basis of bottom-up input (Clark, 2013; Friston, 2005; Friston & Frith, 2015). In other words, predictions are akin to models/internal representations of what may come next, which are created and then checked against input in real time. A prediction is confirmed or disconfirmed depending on whether the model and the input match. Applied to language, this means that a comprehender uses the linguistic and non-linguistic information from the sentence or discourse context to build an internal representation of upcoming words (top-down predictions) and checks the match between that representation and the language input as speech unfolds (for reviews of prediction in language comprehension, see Huettig, 2015; Kuperberg & Jaeger, 2016; Van Petten & Luka, 2012). The object of prediction can be the features of a word (e.g., grammatical class, gender, number) or a specific word

https://doi.org/10.1075/bpa.12.05fou

(e.g., the word *kite* in the sentence *The day was breezy so the boy went outside to fly a ___*). When successful, predictively pre-activating information helps one to infer the interlocutor's message, which facilitates communication. However, prediction is often challenging. Context and input are not always optimal and their efficiency may depend on individual resource limitations, such as reading skills (Huettig & Pickering, 2019; Mishra et al., 2012), vocabulary size (Mani & Huettig, 2012), working memory (Huettig & Janse, 2016; Ito et al., 2018; Otten & Van Berkum, 2009) or processing speed (Huettig & Janse, 2016). These factors also partially account for the lesser and more variable ability to predict in a second language (L2), with some studies reporting prediction in L2 and others not (for a review of prediction in L1 and L2, see Kaan, 2014).

One factor that is unique to bilinguals is potential cross-linguistic influence (CLI). CLI refers to the impact that a bilingual's languages can have on each other, which can either impede (negative transfer) or facilitate (positive transfer) language processing (Gass & Selinker, 1992; Jarvis & Pavlenko, 2008; Odlin, 2003). To date, only a handful of studies have directly tested the effect of CLI on prediction. This chapter provides an overview of the effect of CLI on language processing in general to identify the circumstances under which these effects may play a (positive or negative) role in prediction. We then review the few studies that have observed effects of CLI on predictive processing. Given the scarce literature on this topic, drawing strong conclusions concerning the effects CLI may have on prediction mechanisms are undoubtedly premature at this stage. Hence this chapter aims instead to raise awareness of the importance of considering the role of CLI on prediction in L2 processing and to propose directions for future research.

It is important to underline that this review is by no means exhaustive and focuses mainly on comprehension in reading or listening (even though prediction is thought to play a role in production, as discussed in the section *Cross-linguistic influence and prediction-by-production*). Similarly, we will consider CLI in the framework of various models of language prediction; however, our intention is not to present the models in detail. We direct the interested reader to other chapters in this volume. Most studies reported in this chapter involved late bilinguals, referred to herein as L2 learners or L2 speakers, that is, adults who have learned their L2 after childhood (roughly after age 10). Results from early bilinguals will be clearly specified. Finally, we are aware that the impact of CLI on prediction may depend on factors that are difficult to disentangle in L2 processing, such as proficiency, language exposure or individual differences. The impact of these factors on prediction during L2 comprehension have already been intensely examined in Kaan's (2014) review and will not be discussed in this chapter.

Cross-linguistic influence and language processing in L2

Evidence of CLI comes from studies that have used various types of manipulations. For instance, in an early eye-tracking study, Frenck-Mestre and Pynte (1997) compared sentence reading in French-English and English-French bilinguals using syntactic ambiguity resolution. Sentences contained verbs that had different properties in the bilinguals' two languages, as in *Every time the dog obeyed the pretty little girl showed her approval*. In English, this sentence is ambiguous up to the main verb (*showed*) because the subordinate verb *obey* can be either transitive or intransitive, whereas in French the disambiguation occurs on *obey* which is intransitive. Results showed that, despite a momentary localized effect of transfer (as revealed by a higher number of regressions from the disambiguating region and longer reading times during the re-reading of sentences), similar patterns of eye-movements were obtained independently of the bilinguals' L1 when ambiguity was resolved. These findings converge with other studies that claim that L2 speakers make use of L2 lexical-semantic and syntactic information during L2 sentence processing (Hoover & Dwivedi, 1998; Juffs, 1998; Juffs & Harrington, 1996).

To further investigate CLI during sentence processing, studies have looked at whether L2 speakers use the same processing or parsing strategies as native speakers or whether they tend to transfer the strategies they use in their L1 during L2 comprehension. A handful of studies have addressed this question by examining the resolution of temporarily ambiguous structures, notably relative clause attachment ambiguities (Dussias, 2001, 2003; Felser et al., 2003; Fernández, 2003; Frenck-Mestre, 2002; Papadopoulou & Clahsen, 2003). The manipulation stems from the observation that when processing sentences like *Peter fell in love with the daughter of the psychologist who studied in California* (Dussias, 2003) native speakers of some languages, like English, would interpret the relative clause locally (low attachment), whereas native speakers of other languages (in their respective languages), like Spanish, would interpret it non-locally (high attachment; Cuetos & Mitchell, 1988). In other words, English speakers would interpret that the psychologist studied in California, whereas Spanish speakers would understand the daughter studied in California. Studies using different pairs of languages reached distinct conclusions. While some L2 speakers showed similar attachment preferences as native speakers when processing their L2, others transferred the attachment preferences of their L1, and some did not show any preference for either one or the other (Frenck-Mestre, 2002; Papadopoulou & Clahsen, 2003). These differences have been explained by factors such as language proficiency or length of exposure (Frenck-Mestre, 2002), or have been taken as evidence that structure-building processes in L2 are not the same as in L1 (Felser et al., 2003; Papadopoulou & Clahsen, 2003). The latter claim is in line with Clahsen and Felser's Shallow Structure Hypothesis (Clahsen & Felser, 2006b, 2018) which argues that L2 speakers do not use syntactic information as native speakers do during sentence processing but tend to rely more on lexico-semantic and pragmatic cues.

Similar conclusions that L2 speakers do not make use of syntactic structure/information to the same extent as native speakers have been reached when examining bilinguals' parsing of filler-gap constructions, such as in *wh*-questions (e.g., Marinis et al., 2005). Nonetheless, this view has not been supported universally (e.g., Juffs, 2005; Juffs & Harrington, 1995; Williams et al., 2001). The L1 literature suggests that when native speakers encounter a sentence like *Who did the police know (declare) __ killed the pedestrian?* (Dussias & Piñar, 2010), they tend to link the *wh*-element placed at the beginning of the sentence to its subcategorizing head as soon as possible (e.g., Frazier & Clifton, 1989; Pritchett, 1992). Studies have shown that L2 speakers use the same strategies and experience a processing cost, i.e., a higher cognitive effort to process the sentence, like native speakers do. Their processing strategies seem to be independent of whether *wh*-constructions are similar in their L1 (Williams et al., 2001), but may be affected by proficiency level (Dussias & Piñar, 2010) and cognitive resources (Williams, 2006).

Overall, research on the effect of the L1 syntax during online L2 sentence processing has not reached clear-cut conclusions (see Roberts, 2013, for a review). In relation to CLI effects on prediction, the fact that L2 speakers may rely on different cues than native speakers or use parsing strategies from their L1 when processing their L2 may prevent them from efficiently using the information from the sentence/discourse context to build internal representations of upcoming words.

Furthermore, CLI has also been observed on L1 and L2 lexical representations, due to the co-activation of the bilinguals' two languages. Indeed, since lexical access is thought to be non-selective (Dijkstra & van Heuven, 2002), both the L1 and the L2 lexical representations are activated which can lead to competition. This interlingual competition can occur when features (e.g., phonological, syntactic) are similar across languages (e.g., Chambers & Cooke, 2009; Marian & Spivey, 2003; Weber & Cutler, 2004) or when they are incongruent (e.g., Foucart & Frenck-Mestre, 2011; Sabourin & Stowe, 2008). For instance, Chambers and Cooke (2009) used a visual world paradigm to examine lexical competition from phonological competitors. This paradigm consisted of presenting participants with a visual display containing a target word and competitors; participants' eye-movements were recorded while they listened to audio stimuli. The authors presented English-French bilinguals with audio sentences in French, i.e., *Marie va décrire la poule* ('Marie will describe the chicken'), along with a display containing interlingual competitors (e.g., *chicken* ['poule' /pul/ in French] and *pool* /pul/) and unrelated distracters. The eye-movement patterns revealed that the participants temporarily considered the interlingual competitor. The competition was significantly reduced when the word was embedded in a constrained context, i.e., when the semantic context was incompatible with the competitor (e.g., *Marie va nourrir…* ['Marie will feed…']), suggesting that context information facilitates word recognition (for similar results, see, Fitzpatrick & Indefrey, 2009; Lagrou et al, 2012). Competition has also been observed when syntactic features mismatch. For instance, in an event-related brain potential (ERP) study Foucart and Frenck-Mestre (2011) compared gender processing in French native speakers and German (L1)-French (L2) learners and examined whether performance was affected by proficiency. They presented sentences in French that contained gender violations between the determiner and the noun (Experiment 1) and the postposed adjective and the noun (Experiment 2). Half of the nouns shared gender across languages, while the other half had incongruent gender in L1 and L2. The results showed a P600 effect (ERP component associated with syntactic processing) in both groups. However, within the L2 group, participants with lower proficiency did not show sensitivity to violations when gender was incongruent across languages, even though they were able to correctly assign gender in a post-test. These results are in line with other studies showing that bilinguals' two gender systems interact (e.g., Lemhöfer et al., 2008; Paolieri et al., 2020; Sabourin & Stowe, 2008).

Although not usually framed as CLI, the L2 literature has shown that features or rules unique to the L2, and hence absent from the learners' L1, also affect processing. For example, grammatical gender, a feature that is absent in numerous languages (e.g., English) has often been manipulated to investigate whether L2 speakers are able to learn new features. This has been examined both as concerns whether they can acquire enough lexical knowledge to assign the correct gender to nouns in tests without time limits (offline tasks) and at the ultimate point, when, like native speakers, they are able to apply the agreement rule during real-time processing (online tasks). Results from ERP and eye-tracking studies have demonstrated that with enough proficiency and/or exposure, L2 learners whose L1 does not have grammatical gender show sensitivity to gender violations during online processing (e.g., Foucart & Frenck-Mestre, 2012; Gabriele et al., 2013; Gillon-Dowens et al., 2010, 2011; Keating, 2009; Rossi et al., 2014; Tokowicz & MacWhinney, 2005, but see Sabourin & Stowe, 2008). The ability to learn new features and compute agreement during real-time processing seems to be modulated by various factors such as proficiency (for a review, van Hell & Tokowicz, 2010). This ability has been the topic of debate in the L2 literature, with some theories claiming that new features cannot be acquired, and that syntactic information is not as robust in L2 as in L1 (Clahsen & Felser, 2006a, 2006b, 2018; Hawkins & Chan, 1997), and others claiming that L2 speakers can learn and process new features (Hopp, 2010; Schwartz & Sprouse, 1996). Concerning prediction, these results suggest that the presence/absence of features and rules in the L1 may affect the use of specific cues to pre-activate upcoming words.

This brief review about CLI shows that the experience with a previous language, the speaker's L1, has an effect on L2 processing at different levels, from lexical representations to parsing strategies. In relation to prediction, the focus of the chapter, it is likely that the cross-linguistic effects that affect language processing in general similarly affect prediction mechanisms. Indeed, to build internal representations of upcoming words, comprehenders can use predictive cues (e.g., the gender of a determiner to predict a noun) or the information contained in the sentence/discourse context. Either way, from what has been observed in the CLI literature, it is likely that the more the L1 and the L2 overlap with respect to specific predictive cues or parsing strategies, the more successful the prediction (although lexical competition may occur, as mentioned above). As previously stated, only a handful of studies have directly compared different populations according to their linguistic background with the same materials to examine to what extent CLI affects L2 prediction. Next, we review these studies to evaluate the factors that determine whether cross-linguistic effects play a (positive or negative) role in prediction.

Cross-linguistic influence and prediction in L2

Do cross-linguistic differences in features and rules affect L2 prediction?

The question of whether cross-linguistic differences in features and rules affect the use of cues to predict L2 words' features has been investigated at different linguistic levels (e.g., semantic, lexical). For instance, in a recent study, van Bergen and Flecken (2017) examined whether previous linguistic experience with placement verb semantics modulates prediction. The L2 literature suggests that semantic representations largely overlap across the languages of a bilingual (Francis, 2005), and ERP studies have shown that, despite being sometimes weaker and slightly delayed, semantic processing is fairly similar in L1 and L2 (for review, see, Frenck-Mestre et al., 2014; Moreno et al., 2008). In line with these findings, studies have shown that L2 learners use semantic cues, such as verb subcategorization, to generate predictions like native speakers do (Hopp, 2015; van Bergen & Flecken, 2017), even if semantic activation during prediction may be slower and weaker than in the L1 (Dijkgraaf et al., 2019). Nevertheless, prediction at the semantic level also seems to be affected by CLI, as revealed by van Bergen and Flecken's (2017) eye-tracking study. Using a visual world paradigm, they examined L1 (Dutch) and L2 (German, English and French native speakers, learners of L2 Dutch) participants' anticipatory patterns of fixations to objects while they listened to sentences. Crucially, the placement verb semantics differs across languages; while Dutch and German use different verbs depending on the position of an object on a surface (e.g., *zetten* 'put. STAND' and *leggen* 'put.LIE', in Dutch), English (*to put*) and French (*mettre*) do not specify the object position. The displays contained a target object compatible only with the semantics of the Dutch verb; a competitor object compatible with the semantics of the verb's translation equivalents in English and French, and two filler objects. The pattern of fixations revealed that German native speakers listening to Dutch anticipated the object as soon as they heard the verb, akin to Dutch natives. On the other hand, French and English participants did not show anticipatory fixations. These findings suggest that semantic similarity across L1 and L2 facilitates prediction during sentence comprehension in L2. They support the idea that comprehenders who have previous experience with specific predictive cues in their L1 will use them for prediction in their L2 more easily than comprehenders who have no previous experience with these cues.

The effect of CLI on L2 prediction has also been examined in the case of lexical and syntactic differences in L1 and L2 grammatical gender. Hopp and Lemmerth (2018) compared predictive gender processing in German native speakers and Russian-German L2 speakers using a visual-world paradigm. German and Russian both have a three-gender system but the same noun may not necessarily share the same gender across languages. Moreover, while gender is marked on suffixes in both languages for adjectives, for the articles, the syntactic agreement rule differs (gender is syntactically realized on postnominal suffixes in Russian but on prenominal articles in German). Participants were presented with displays containing four objects and their eye-movements were recorded while they listened to sentences containing an article (e.g., *Wo ist der/die/das gelbe [N]?* 'Where is the MASC/FEM/NEUT yellow [N]?') or an adjective (e.g., *Wo ist ein kleiner/s gelber/s [N]?* 'Where is a small MASC/NEUT yellow [N]?'). Results showed a native-like pattern for advanced L2 speakers in both conditions. However, in high-intermediate learners, similar patterns were observed when the syntactic rule was similar across languages (adjectives), but when the rule was different similar patterns emerged only when nouns shared gender. These findings confirm that cross-linguistic differences affect the use of predictive cues in L2, but also suggest that the L1 influence is reduced as proficiency increases.

Can L2 speakers use features that do not exist in their L1 as prediction cues?

As previously mentioned, cross-linguistic differences do not only refer to the conflict of an existing feature across languages but also to the absence of the L2 feature in the L1, which seems to create difficulties when it comes to predicting. For instance, in a visual-world paradigm study, Lew-Williams and Fernald (2010) showed that late English-Spanish bilinguals, who do not have grammatical gender in their native language, did not use gender-marked articles to predict target nouns, unlike Spanish native speakers. Hopp (2015) showed similar results with case. In a visual-world eye-tracking experiment, he demonstrated that, while German native speakers integrate both morpho-syntactic (i.e., case marking) and lexico-semantic information (i.e., verb semantics) to generate predictions, English-German L2 learners only used lexico-semantic information. This observation was true independently of the learners' proficiency. Along the same lines, in a visual-world paradigm study, Mitsugi and MacWhinney (2016) investigated whether English-Japanese L2 learners use case markers to predict upcoming words when processing sentences containing either the monotransitive or ditransitive constructions. Like Hopp, they reported that native speakers (Japanese) used syntactic cues to predict upcoming auditory words but L2 learners (English-Japanese) did not. The effect of L1 and L2 overlap was also observed in L2 Korean, with the learners whose L1 has case marking (Kazakh) showing similar prediction patterns as native speakers, and learners without case marking in their L1 (French) did not (Frenck-Mestre et al., 2019). These findings suggest that L2 speakers have difficulty using syntactic cues to predict elements when the feature does not exist in their L1.

However, other studies have shown that high-proficient L2 speakers use these cues in the same way as native speakers. For instance, in an eye-tracking study, Dussias et al. (2013) examined whether high- and low-proficient L2 speakers of an L1 that does not have gender (English) and one that does (Italian) use gender as a predictive cue in their L2 Spanish. They presented participants with sentences containing target items preceded by an article that agreed in gender with two pictures displayed on a screen or agreed with only one of them. Native speakers looked at the target picture sooner on different-gender trials than on same-gender trials. Italian L2 speakers revealed a similar predictive pattern (but only for feminine words), which suggests positive CLI. Highly proficient English L2 speakers used gender as a predictive cue, like native speakers, whereas low-proficient L2 speakers did not. The authors concluded that both cross-linguistic similarity and proficiency affect the ability to use morphosyntactic predictive cues. Similarly, Hopp (2013) showed that L2 speakers that have a good command of their L2 gender system in production are better at using these cues during real-time comprehension compared to those who show variability in their production. The same author later demonstrated that, indeed, with previous training on gender assignment aiming at reducing lexical variability, English learners of German showed predictive gender processing (Hopp, 2016). These findings suggest that the difficulty in using morphosyntactic cues online to predict upcoming words due to the absence of a feature in the L1 can be overcome with sufficient proficiency (but see, Grüter et al., 2012 and Hopp, 2015).

Is the use of the sentence context in L2 prediction affected by CLI?

So far, we have focused on L2 speakers' ability to use specific predictive cues, e.g., gender, to pre-activate the features of upcoming words and how this ability may be modulated by CLI. However, prediction can also be based on a combination of multiple sources of information contained in the sentence context. This information must be integrated incrementally to serve as prediction cues. In this case, the multiple sources of information contained in the sentence context are integrated incrementally to serve as prediction cues. Studies investigating sentence processing in general suggest that L2 speakers do not always rely on the same cues as native speakers when processing sentences in real time or that they use parsing strategies from their L1 (Clahsen & Felser, 2006a, 2006b, 2018; Foucart et al., 2015; Pan et al., 2015). Hence, it is not clear whether L2 speakers predict the same level of detail based on the meaning of the sentence as native speakers. Indeed, prediction requires even more resources than incremental parsing given that the internal representation must be built before the input is available for integration (Hopp, 2015). It is important to note that the comprehender´s level of engagement in prediction has been argued to depend on whether it is valuable for comprehension and that it may not even be necessary for comprehension (Huettig, 2015; Huettig & Mani, 2016). The necessity of prediction in comprehension is particularly relevant in L2. Indeed, if prediction is too costly, L2 speakers may not even try to predict, especially in cases where CLI makes processing even more effortful. Various studies have shown that L2 speakers can integrate information to predict but mostly when the context is optimal (e.g., highly constrained sentences and similar features, as described below).

For instance, in two ERP studies, Foucart et al. (2014, sentence reading; 2016 sentence listening) used a similar design as in Wicha and collaborators' study (Wicha et al., 2004), and presented Spanish native speakers, French-Spanish late bilinguals, and Spanish-Catalan early bilinguals with high-constraint sentences in Spanish. The critical noun was either expected or not (*El pirata tenía el mapa secreto, pero nunca encontró el tesoro* [masc]/*la gruta* [fem] *que buscaba*. 'The pirate had the secret map, but he never found the treasure/the cave he was looking for.'), and crucially, the expected and unexpected nouns were of opposite gender in Spanish. ERPs, in particular, an N400 component at the critical noun (an ERP component, usually associated with lexico-semantic integration), suggested that the integration of an unexpected noun (*gruta*) was costlier than that of an expected noun (*tesoro*) in the three groups. Interestingly, an N400 effect on the preceding article also revealed extra processing effort when the gender was incongruent with the gender of the expected noun in all three groups. The results showed that L2 speakers integrate information from prior context in real time (including their world knowledge) to predict upcoming nouns and use gender cues to maintain their prediction (i.e., they verify that the gender of the article preceding the expected noun matches that of the noun to confirm their prediction). However, although the findings suggest that L2 speakers can predict like native speakers, it is important to note that the expected noun was selected so that its gender was always congruent across languages, and that the agreement rule between the article and the noun was similar in all three languages. Hence, the experimental conditions involved positive CLI and were, therefore, optimal to favor prediction.

Other studies, however, have shown that L2 speakers' ability to predict based on the preceding linguistic context may be negatively affected by CLI. For instance, Martin and collaborators (2013) replicated DeLong et al.'s (2005) study, which exploited the phonological rule in English according to which the indefinite article *a* becomes *an* when preceding a vowel. The critical noun in high-constraint sentences like *The day was breezy so the boy went outside to fly a kite* was manipulated so that it was either expected (*a kite*) or unexpected (*an airplane*). The original ERP patterns from DeLong et al. (2005) showed extra processing not only on unexpected critical nouns but also on their preceding articles, suggesting that readers pre-activate the phonological form of predicted words (but see Nieuwland et al., 2018). Martin et al. (2013) presented similar sentences to English native speakers and late Spanish–English bilinguals. When reading the sentence *the boy went outside to fly a*…, the ERPs time-locked on the noun showed a significant difference between the expected and the unexpected noun in both L1 and L2 speakers. However, while a difference was also observed when native speakers expected the article *a* but read the article *an*, indicating they had predicted the word *kite*, such a difference was absent in L2. The authors concluded that lexical prediction is weaker in L2. Interestingly, the phonological rule of the article was unique to the L2. Hence, although L2 speakers showed good knowledge of the rule offline, it is possible that they predicted the critical word from the context but that they did not apply the rule online or they did so not fast enough for it to be visible in the ERP data on the article.

Another potential explanation for the reduced prediction observed by the Martin et al. (2013) is that L2 speakers may try to predict in their two languages simultaneously, and therefore may suffer interference from the L1 word. This account was investigated in a visual-world paradigm study by Ito, Pickering and Corley (2018). They presented English native speakers and Japanese-English bilinguals with high-constraint sentences spoken in English (e.g., *The tourists expected rain when the sun went behind the …*) and a visual display with objects corresponding to the target word (*cloud*; Japanese: *kumo*), an English phonological competitor (*clown; piero*), a Japanese competitor (*bear; kuma*), and an unrelated object. Both

L1 and L2 speakers looked predictively at the target object, but L2 speakers were slower. Native English participants looked predictively at the English phonological competitor but L2 participants did not do so, suggesting that bilinguals did not predict phonological information, in line with Martin et al. (2013). L1 activation during L2 comprehension did not occur or it was too weak to generate competition. Hence, CLI did not seem to explain why L2 speakers do not predict phonological information to the same extent as native speakers. Note that to account for the absence of phonological competitor effects in their study and the strong effect found in previous work (e.g., Chambers & Cooke, 2009), the authors evoked the role of participants' proficiency. This was supported by the observation that participants who had more L2 English exposure looked more at the English phonological competitor.

In relation to language exposure, the activation of the L1 during L2 prediction has been observed even with early bilinguals. Molinaro et al. (2017) tested Basque-Spanish and Spanish-Basque balanced bilinguals in an ERP study to investigate whether linguistic experience modulates the use of predictive cues. Participants were presented with high-constraint sentences in which the gender of the (un-) expected critical noun was manipulated (similar design as in Foucart et al., 2014). The ending of the noun was either transparent or opaque. Crucially, the Basque language does not have grammatical gender; noun endings (post-nominal suffixes) are particularly relevant in this language. Event-related potentials (N200-N400) and oscillatory activity in the low beta-band (15–17 Hz) for the critical noun and the preceding article revealed that both groups predicted the critical word, but participants who had learned Basque first also displayed visual word form predictions for transparent words, in accordance with the relevance of noun endings in their L1. The authors concluded that prediction mechanisms are shaped by early language exposure and that bilinguals apply the regularities they have learned from their L1 when predicting in L2.

Overall, the findings reported above suggest that L2 speakers can integrate information from the prior sentence context to predict upcoming words, at least when the context is optimal (e.g., high-constraint sentences). Although CLI has been observed in early bilinguals, it seems that L1 lexical activation during L2 comprehension in late L2 learners may be too weak to generate competition predictively (at least for phonological information). Little is known, however, about L2 speakers' ability to predict information from the wider discourse context, which requires even more cognitive resources.

Is the use of the discourse context in L2 prediction affected by CLI?: First steps

Discourse processing requires the integration of multiple types of information, such as lexico-semantic and syntactic information, but also prior information from the context (discourse-pragmatic information). It has been argued that L2 speakers seem to be less sensitive to discourse-level properties than native speakers (Hopp, 2009). Discourse processing also often involves establishing coreference and inferences. Although late bilinguals have been shown to be able to make inferences and find the appropriate reference of pronouns, they seem to have more difficulties in updating information than native speakers (Foucart et al., 2016; Kohlstedt & Mani, 2018; Roberts et al., 2008). Hence, it is likely that using cues from the discourse context to predict upcoming words may also be more difficult for L2 speakers. To our knowledge, very few studies have addressed this question, and when they did, it was not always their main research question.

For instance, Grüter et al. (2017) used a written story continuation task to examine the role of two discourse-level factors, event structure and referential form, in L2 speakers' referential choices. Japanese- and Korean-L1 learners of English were asked to continue (im-)perfective sentences (*Emily brought/was bringing a drink to Melissa. (She)…*) either in a pronoun prompt or a free-prompt condition. Discourse-level biases were similar in native and non-native speakers when they involved referential form, however, L2 speakers' referential choices were not affected by event structure as much as native speakers'. Although the study was not designed to directly test prediction in L2, Japanese and Korean were chosen as L1 because the potentially predictive cue – grammatical aspect marking – is implemented similarly in these language as in English, which favored a potential positive CLI. The authors propose an account of the results based on the L2 prediction literature and suggested that both L1 and L2 speakers use grammatical aspect to build event structures, but that L1 speakers use the event structure proactively to predict a coherence relation and a subject referent before the completion of the context sentence. On the other hand, they posited that L2 speakers do not incrementally predict and rather wait until the end of the context to build a coherence relation.

More recently, Kim and Grüter (2020) conducted a visual-world paradigm study to examine the extent to which L2 learners use implicit causality information to create internal representations of upcoming reference during real-time comprehension. Implicit causality refers to the biases that certain verbs generate towards either their subject or object in a causal dependent clause. Native speakers of English and Korean learners of English were presented with sentences that contained a bias towards either the first or second noun phrase of a clause while viewing two human faces. The critical sentence was preceded by a context such as: *Nathan and*

Owen used to study together in the library (context), *Nathan disturbed Owen all the time because he needed some help with his homework* (critical). The ending was either bias consistent or inconsistent. Eye-tracking patterns revealed an early use of implicit causality information in both L1 and L2 listeners, however; the effect was both weaker and delayed in the L2 group. This result was not modulated by the learners' proficiency. The findings suggest that L2 speakers can use the information from the discourse context to generate predictions, however, they do so to a lesser extent than native speakers. In addition to the difficulty in integrating information online to incrementally update the discourse models, the activation of lexical representations for the verbs may be weaker in L2 than in L1.

Although these studies were not directly designed to investigate the use of the discourse context in L2 prediction, they provide preliminary information as they suggest that L2 speakers can incrementally use cues from the discourse context to predict upcoming words but to a limited extent. The next step would be to examine whether CLI contributes to L2 speakers' limited predictive ability. For instance, a comprehender may have more difficulty establishing coreference between a pronoun and a noun when the pronoun refers to a noun that does not share the same gender in the L1 and the L2 compared to when it does. Further research is needed.

Cross-linguistic influence and models of language prediction

Models of language prediction have associated prediction with processing, production, prediction errors, priming, implicit learning and adaptation (Chang et al., 2006; Dell & Chang, 2014; Huettig, 2015; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Pickering & Garrod, 2007, 2013). We will examine the effect CLI can have on the various mechanisms underlying prediction proposed by these models.

Cross-linguistic influence and error-based implicit learning

One theory that has been put forward is that prediction allows adaptation and implicit learning (Chang et al., 2006; Dell & Chang, 2014; Jaeger & Snider, 2013). The idea is that, when communicating, comprehenders use incremental information to predict upcoming linguistic events and when the bottom-up input does not match their internal representations, this leads to a prediction error. When an error occurs, comprehenders adjust their predictions to better match their interlocutor's speech in the future. In other words, comprehenders adapt their comprehension and production to align with their interlocutors' productions by using previously used syntactic structures or words, for example. This alignment is referred to as priming. This priming is thought to be generated by the degree of mismatch between the prediction and the input, with larger degrees of mismatch leading to larger adjustments of the prior knowledge.

The observation that priming effects are larger for low-frequency structures is referred to as the inverse frequency effect (Bernolet & Hartsuiker, 2010). L2 learners usually show similar priming effects as native speakers (for a review, see Jackson, 2018), hence, they should, in theory, benefit from prediction errors as much as native speakers do. Even more so, the logic would expect that, given that a structure is generally less frequent in L2 than in L1 (due to language experience), the inverse frequency effect should be even larger in L2 speakers, who should then benefit even more from an error-based implicit learning mechanism. Following this logic, cross-linguistic differences should boost learning (Flett et al., 2013; Jackson & Ruf, 2017; Kaan & Chun, 2018). If we take the example of grammatical gender, if an L2 learner encounters an article related to a noun that has incongruent gender in the two languages (e.g., feminine in L1 but masculine in L2) they may experience a prediction error because the input would not match the internal representation of the L1 feminine article corresponding to the expected noun in L1. The learner would then have to adapt to the interlocutor and learn that the noun in their L2 is masculine. However, for a prediction error to occur, prediction must first take place. For this reason, it has been suggested that late bilinguals' reduced ability to predict (Grüter et al., 2014; Kaan, 2014; Martin et al., 2013) may diminish the impact of prediction errors and, consequently, the possibility of learning from adaptation (Hopp, this volume; Phillips & Ehrenhofer, 2015). Moreover, cross-linguistic differences may also have the reverse effect, hence, instead of experiencing a prediction error that would lead to learning, they would stop predicting. For instance, if gender largely overlaps across a speaker's two languages, L2 speakers would successfully use gender cues to predict. However, if learners encounter words of different gender across languages, then gender cues may become less stable and prediction would stop. This pattern was suggested in Hopp's (2016) study in which German native speakers stopped using the gender cues due to large lexical variability in gender agreement (but see Dussias et al., 2013; Johnson Fowler & Jackson, 2017). This assumption is consistent with the claim that prediction may depend on its expected utility (Kuperberg & Jaeger, 2016). Thus, whether L2 speakers can benefit from prediction errors to learn is not clear (see Grüter et al., this volume; Hopp, this volume). Recent studies have suggested that the difference between L1 and L2 speakers may not lie in their ability to predict but rather in their use of prediction errors (Jackson & Hopp, 2020; Kaan et al., 2019), which leads to implicit learning in L1, but may only trigger short-term priming in L2. The results from Hopp and Lemmerth (2018), reviewed earlier, converge with this hypothesis since high-intermediate Russian-German L2 learners showed similar predictive ability as native speakers when the syntactic rule was similar across language (adjectives), but only when nouns were lexically congruent when the rule was different (article). These results suggest that L2 learners were able to predict but did not successfully use prediction error for learning when nouns were incongruent. Thus, although recent findings suggest that L2 speakers may not benefit from prediction errors to the same extent as native speakers, further research is still needed to confirm this hypothesis and to examine whether CLI modulates the use of prediction errors.

Cross-linguistic influence and prediction-by-production

In some approaches, prediction is thought to be production-based, which means that the same mechanism used by comprehenders to generate prediction is used during production (Dell & Chang, 2014; Ito & Pickering, this volume; Pickering & Gambi, 2018; Pickering & Garrod, 2007, 2013). The argument is that the comprehender uses various types of information coming from the linguistic context as well as from extra-linguistic information to infer the interlocutor's message. This inference then goes through the comprehender's production system, which allows retrieving and generating production representations. These representations correspond to the prediction of what the interlocutor is about to say. According to Pickering and Gambi (2018), prediction-by-production is a mechanism that occurs at any linguistic level, and that, although it helps, it is not mandatory for comprehension (see also Ito & Pickering, this volume). Chang et al. (2006) proposed that production skills develop with the ability to predict. The essential role of production for prediction has recently been demonstrated in a study by Martin et al. (2018) in which prediction was impeded when the production system was taxed in the sentence context. There are two relevant points for L2 prediction. First, this prediction-by-production theory implies a causality between production and prediction; the better the production skills, the better the prediction (Rommers et al., 2015). This point was supported by Hopp's (2013, 2016) studies in which L2 speakers that had a good command of their L2 gender system in production and reduced lexical variability were better at using these cues during real-time comprehension compared to those who showed variability in their production. Second, the whole prediction-by-production process must happen before the interlocutor starts producing the expected word. Given that production in L2 is usually not as fluent as in L1 and that processing is usually slower, the additional presence of cross-linguistic differences may compromise prediction. The difficulty may be even more salient when predictive cues require processing morphosyntactic information (compared to semantic or lexical information) since comprehension usually precedes production in L2 acquisition of morphosyntax (Lardiere, 2000). This is begging the question why production is harder. Ito & Pickering (this volume) speculate that semantic prediction can take place through association. In the production process, morphosyntactic information is planned later than semantic information, and phonological information even later still. Under time or resource constraints, prediction at the morpho-syntactic or even phonological level may not occur. Let us take the example of grammatical gender again. Production of gender markings in L2 has been shown to be affected by the absence of gender in the L1 or by a conflict across languages (Foucart, 2008; Grüter et al., 2012; Hopp, 2013), due to lexical competition, as for comprehension, as mentioned earlier. Thus, if a Russian-German learner uses her/his production representations (e.g., feminine article in L1) to predict what a German native speaker will say, there may be greater surprisal when they hear the interlocutor's output (e.g., masculine article in L2). As suggested earlier, this surprisal could either be beneficial (implicit learning) or could make learners stop using unstable gender cues. Interestingly, the production-comprehension cross-linguistic conflict could go both ways. Indeed, it is likely that if native speakers base their prediction on their production, when they hear an incorrect input produced by a non-native speaker, they will most likely stop using the unstable cue to predict. This is consistent with studies that have shown that prediction mechanisms are affected by hesitations (i.e., comprehenders stop predicting when sentences contain hesitations like 'em…' which render the context less reliable) or foreign accent speech (Corley et al., 2007; Romero-Rivas et al., 2016).

Conclusions and directions for future research

After providing a background on cross-linguistic influence (CLI) during language processing in general to identify the circumstances under which it may play a (positive or negative) role in prediction, this chapter reviewed the relatively few studies that have addressed this question to date. As mentioned at the outset, the literature is still scarce and drawing strong conclusions is premature. Nevertheless, the findings reviewed here show that, when cues are similar across languages and/or when the context is optimal, prediction usually occurs in L2. Hence, it seems that prediction mechanisms are not fundamentally different in native and non-native speakers (see, Kaan, 2014 for discussion). Differences observed in L1 and L2 may originate from L1 and L2 users' level of engagement as suggested in approaches like the Reduced Ability to Generate Expectations (Grüter et al., 2014) that put forward that L2 speakers have a reduced ability to predict because they do not engage in proactive processing to the same extent as native speakers and integrate information instead. While previous experience with specific predictive cues seems to facilitate prediction, it can also generate competition between the L1 and L2 features and rules, however, future research is needed to determine whether prediction at semantic, phonological and lexical levels are similarly affected by CLI. Another question that will need to be addressed is whether the absence of lexical competition is due to weaker lexical activation in L2 or to the fact that lexical representations are not pre-activated by the incremental processing of the prior context. Moreover, to date, most studies have examined CLI on predictive cues, but whether or not CLI makes the sentence context too unreliable to extract information to generate predictions of upcoming words needs to be examined (e.g., if the presence of a word that has different gender in L1 and L2 within the sentence context reduces lexical prediction of upcoming words). Similarly, the contribution of CLI to L2 speakers' limited predictive ability when having to integrate multiple types of information, including discourse-pragmatic information, still needs to be investigated.

Naturally, results showing no or reduced prediction in L2 can be accounted for by many other factors than cross-linguistic differences that were not discussed in this chapter, such as cognitive resources (Juffs & Harrington, 2011), cue-weighting (Bates & MacWhinney, 1989) or automatic processing (Ullman, 2005). In line with theories of L2 processing (Dekydtspotter et al., 2006; Hopp, 2010; McDonald, 2006), we assume that once proficiency increases and cognitive resources are less taxed during L2 processing, the impact of cross-linguistic differences on prediction mechanisms may be reduced, resulting in native-like prediction in L2. Further research is needed to investigate all these unanswered questions.

References

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In E. Bates & B. MacWhinney (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–76). Cambridge University Press.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition*, 114(3), 455–461. https://doi.org/10.1016/j.cognition.2009.11.005

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 35(4), 1029–1040. https://doi.org/10.1037/a0015901

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, 27(1), 107–126. https://doi.org/10.1017/S0142716406060206

Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, 10(12), 564–570. https://doi.org/10.1016/j.tics.2006.10.002

Clahsen, H., & Felser, C. (2018). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition*, 40(3), 693–706. https://doi.org/10.1017/S0272263117000250

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: hesitations in speech affect language comprehension. *Cognition*, 105(3), 658–668. https://doi.org/10.1016/j.cognition.2006.10.010

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30(1), 73–105. https://doi.org/10.1016/0010-0277(88)90004-2

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. Grantham O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 33–40). Cascadilla Proceedings Project.

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634). https://doi.org/10.1098/rstb.2012.0394

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121. https://doi.org/10.1038/nn1504

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dussias, P. E. (2001). Sentence parsing in fluent Spanish–English bilinguals. In J. L. Nicol (Ed.), *One mind, two languages* (pp. 159–176). Basil Blackwell.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1. *Studies in Second Language Acquisition*, 25(4), 529–557. https://doi.org/10.1017/S0272263103000238

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English second language speakers. *Second Language Research*, 26(4), 443–472. https://doi.org/10.1177/0267658310373326

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, 24(3), 453–489. https://doi.org/10.1017/S0142716403000237

Fernández, E. M. (2003). *Bilingual sentence processing: Relative clause attachment in English and Spanish*. John Benjamins. https://doi.org/10.1075/lald.29

Fitzpatrick, I., & Indefrey, P. (2009). Lexical competition in nonnative speech comprehension. *Journal of Cognitive Neuroscience*, 22(6), 1165–1178. https://doi.org/10.1162/jocn.2009.21301

Flett, S., Branigan, H. P., & Pickering, M. J. (2013). Are non-native structural preferences affected by native language preferences? *Bilingualism: Language and Cognition*, 16(4), 751–760. https://doi.org/10.1017/S1366728912000594

Foucart, A. (2008). Grammatical gender processing in French as a first and a second language (Doctoral dissertation). Université Aix-Marseille I/University of Edinburgh. Edinburgh Research Archive.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1-L2 syntactic similarity. *Bilingualism: Language and Cognition*, 14(3), 379–399. https://doi.org/10.1017/S136672891000012X

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, 66(1), 226–248. https://doi.org/10.1016/j.jml.2011.07.007

Foucart, A., Garcia, X., Ayguasanosa, M., Thierry, G., Martin, C. D., & Costa, A. (2015). Does the speaker matter? Online processing of semantic and pragmatic information in L2 speech comprehension. *Neuropsychologia*, 75, 291–303. https://doi.org/10.1016/j.neuropsychologia.2015.06.027

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1461–1469. https://doi.org/10.1037/a0036756

Foucart, A., Romero-Rivas, C., Gort, B. L., & Costa, A. (2016). Discourse comprehension in L2: Making sense of what is not explicitly said. *Brain and Language*, 163, 32–41. https://doi.org/10.1016/j.bandl.2016.09.001

Foucart, A., Ruiz-Tada, E., & Costa, A. (2016). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 19(1). https://doi.org/10.1017/S1366728915000486

Francis, W. S. (2005). Bilingual semantic and conceptual representation. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 251–267). Oxford University Press.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4(2), 93–126. https://doi.org/10.1080/01690968908406359

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). North Holland. https://doi.org/10.1016/S0166-4115(02)80012-7

Frenck-Mestre, C., Kim, S. K., Choo, H., Ghio, A., Herschensohn, J., & Koh, S. (2019). Look and listen! The online processing of Korean case by native and non-native speakers. *Language, Cognition and Neuroscience*, 34(3), 385–404. https://doi.org/10.1080/23273798.2018.1549332

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, 50A(1), 119–148. https://doi.org/10.1080/027249897392251

Frenck-Mestre, C., Sneed German, E., & Foucart, A. (2014). Qualitative differences in native and non-native semantic processing as revealed by ERPs. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of bilingual memory* (pp. 237–255). Springer. https://doi.org/10.1007/978-1-4614-9218-4_12

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K., & Frith, C. (2015). Active inference, communication and hermeneutics. *Cortex*, 68, 129–143. https://doi.org/10.1016/j.cortex.2015.03.025

Gabriele, A., Fiorentino, R., & Bañón, J. A. (2013). Examining second language development using event-related potentials: A cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism*, 3(2), 213–232. https://doi.org/10.1075/lab.3.2.04gab
Gass, S., & Selinker, L. (1992). *Language transfer in language learning*. John Benjamins. https://doi.org/10.1075/lald.5
Gillon-Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish – evidence from event related potentials. *Neuropsychologia*, 49(7), 1651–1659. https://doi.org/10.1016/j.neuropsychologia.2011.02.034
Gillon-Dowens, M., Vergara, M., Barber, H., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, 22(8), 1870–1887. https://doi.org/10.1162/jocn.2009.21304
Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191–215. https://doi.org/10.1177/0267658312437990
Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2. *Linguistic Approaches to Bilingualism*, 7(2), 199–229. https://doi.org/10.1075/lab.15011.gru
Grüter, T., Rohde, H., & Schafer, A. J. (2014). The role of discourse-level expectations in non-native speakers' referential choices. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the 38th Annual Boston University Conference on Language Development* (pp. 179–191). Cascadilla Press.
Hawkins, R., & Chan, C. Y. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis.' *Second Language Research*, 13(3), 187–226. https://doi.org/10.1191/026765897671476153
Hoover, M., & Dwivedi, V. (1998). Syntactic processing by skilled bilinguals. *Language Learning*, 48(1), 1–29. https://doi.org/10.1111/1467-9922.00031
Hopp, H. (2009). The syntax discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition*, 12(4), 463–483. https://doi.org/10.1017/S1366728909990253
Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120(4), 901–931. https://doi.org/10.1016/j.lingua.2009.06.004
Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. https://doi.org/10.1177/0267658312461803
Hopp, H. (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014
Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32(2), 277–307. https://doi.org/10.1177/0267658315624960
Hopp, H., & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, 40(1), 171–199. https://doi.org/10.1017/S0272263116000437
Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, 31(1), 19–31. https://doi.org/10.1080/23273798.2015.1072223

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Sciences*, 23(6), 464–475. https://doi.org/10.1016/j.tics.2019.03.008

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, 98, 1–11. https://doi.org/10.1016/j.jml.2017.09.002

Jackson, C. N. (2018). Second language structural priming: A critical review and directions for future research. *Second Language Research*, 34(4), 539–552. https://doi.org/10.1177/0267658317746207

Jackson, C. N., & Hopp, H. (2020). Prediction error and implicit learning in L1 and L2 syntactic priming. *International Journal of Bilingualism*, 24(5–6), 895–911. https://doi.org/10.1177/1367006920902851

Jackson, C. N., & Ruf, H. T. (2017). The priming of word order in second language German. *Applied Psycholinguistics*, 38(2), 315–345. https://doi.org/10.1017/S0142716416000205

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Jarvis, S., & Pavlenko, A. (2008). *Cross-linguistic influence in language and cognition*. Routledge. https://doi.org/10.4324/9780203935927

Johnson Fowler, C., & Jackson, C. N. (2017). Facilitating morphosyntactic and semantic prediction among second language speakers of German. *Journal of Cognitive Psychology*, 29(8), 883–901. https://doi.org/10.1080/20445911.2017.1353517

Juffs, A. (1998). Main verb versus reduced relative clause ambiguity resolution in L2 sentence processing. *Language Learning*, 48(1), 107–147. https://doi.org/10.1111/1467-9922.00034

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, 21(2), 121–151. https://doi.org/10.1191/0267658305sr255oa

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition*, 17(4), 483–516. https://doi.org/10.1017/S027226310001442X

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language. *Sentence Processing. Language Learning*, 46(2), 283–323. https://doi.org/10.1111/j.1467-1770.1996.tb01237.x

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44(02), 137–166. https://doi.org/10.1017/S0261444810000509

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E., & Chun, E. (2018). Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, 21(2), 228–242. https://doi.org/10.1017/S1366728916001231

Kaan, E., Futch, C., Fernández Fuertes, R., Mujcinovic, S., & Álvarez De La Fuente, E. (2019). Adaptation to syntactic structures in native and nonnative sentence comprehension. *Applied Psycholinguistics*, 40(1), 3–27. https://doi.org/10.1017/S0142716418000437

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, 59(3), 503–535. https://doi.org/10.1111/j.1467-9922.2009.00516.x

Kim, H., & Grüter, T. (2020). Predictive processing of implicit causality in a second language a visual-world eye-tracking study. *Studies in Second Language Acquisition*, Advance online publication. https://doi.org/10.1017/S0272263120000443

Kohlstedt, T., & Mani, N. (2018). The influence of increasing discourse context on L1 and L2 spoken language processing. *Bilingualism: Language and Cognition*, 21(1), 121–136. https://doi.org/10.1017/S1366728916001139

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2012). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition*, 16(3), 508–517. https://doi.org/10.1017/S1366728912000508

Lardiere, D. (2000). Mapping features to forms in SLA. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 102–129). Blackwell.

Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, 59(3), 312–330. https://doi.org/10.1016/j.jml.2008.06.005

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843–847. https://doi.org/10.1037/a0029284

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, 6(2), 97–115. https://doi.org/10.1017/S1366728903001068

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27(1), 53–78. https://doi.org/10.1017/S0272263105050035

Martin, C. D., Branzi, F. M., & Bar, M. (2018). Prediction is production: The missing link between language production and comprehension. *Scientific Reports*, 8(1), 1–9. https://doi.org/10.1038/s41598-018-19499-4

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1), 1–10. https://doi.org/10.16910/jemr.5.1.3

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881

Molinaro, N., Giannelli, F., Caffarra, S., & Martin, C. D. (2017). Hierarchical levels of representation in language prediction: The influence of first language acquisition in highly proficient bilinguals. *Cognition*, 164, 61–73. https://doi.org/10.1016/j.cognition.2017.03.012

Moreno, E. M., Rodriguez-Fornells, A., & Laine, M. (2008). Event-Related Potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, 21(6), 477–508. https://doi.org/10.1016/j.jneuroling.2008.01.003

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, 7. https://doi.org/10.7554/eLife.33468

Odlin, T. (2003). Cross-linguistic Influence. In C. Doughty & M. Long (Eds.), *Handbook on second language acquisition* (pp. 436–486). Blackwell. https://doi.org/10.1002/9780470756492.ch15

Otten, M., & van Berkum, J. J. A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, 1291, 92–101. https://doi.org/10.1016/j.brainres.2009.07.042

Pan, H.-Y., Schimke, S., & Felser, C. (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism*, 19(3), 298–313. https://doi.org/10.1177/1367006913515769

Paolieri, D., Demestre, J., Guasch, M., Bajo, T., & Ferré, P. (2020). The gender congruency effect in Catalan-Spanish bilinguals: Behavioral and electrophysiological evidence. *Bilingualism: Language and Cognition*, 23(5), 1045–1055. https://doi.org/10.1017/S1366728920000073

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, 25(4), 501–528. https://doi.org/10.1017/S0272263103000214

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, 5(4), 409–453. https://doi.org/10.1075/lab.5.4.01phi

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3), 105–110. https://doi.org/10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. The University of Chicago Press.

Roberts, L. (2013). Sentence processing in bilinguals. In R. van Gompel (Ed.), *Sentence processing: Current issues in language*. (pp. 221–246). Psychology Press.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse. *Studies in Second Language Acquisition*, 30(3), 333–357. https://doi.org/10.1017/S0272263108080480

Romero-Rivas, C., Martin, C. D., & Costa, A. (2016). Foreign-accented speech modulates linguistic anticipatory processes. *Neuropsychologia*, 85, 245–255. https://doi.org/10.1016/j.neuropsychologia.2016.03.022

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, and Psychophysics*, 77(3), 720–730. https://doi.org/10.3758/s13414-015-0873-x

Rossi, E., Kroll, J. F., & Dussias, P. E. (2014). Clitic pronouns reveal the time course of processing gender and number in a second language. *Neuropsychologia*, 62(1), 11–25. https://doi.org/10.1016/j.neuropsychologia.2014.07.002

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are L1 and L2 processed similarly. *Second Language Research*, 24(3), 397–430. https://doi.org/10.1177/0267658308090186

Schwartz, B., & Sprouse, R. (1996). L2 Cognitive states and the full transfer/full access model. *Second Language Research*, 12(1), 40–72. https://doi.org/10.1177/026765839601200103

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An Event-Related Potential investigation. *Studies in Second Language Acquisition*, 27(2), 173–204. https://doi.org/10.1017/S0272263105050102

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice*. (pp. 141–178). Georgetown University Press.

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, 26(1), 43–74. https://doi.org/10.1177/0267658309337637

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50(1), 1–25. https://doi.org/10.1016/S0749-596X(03)00105-0

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, 16(7), 1272–1288. https://doi.org/10.1162/0898929041920487

Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, 9(1), 71–88. https://doi.org/10.1017/S1366728905002385

Williams, J. N., Möbius, P., & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics*, 22(4), 509–540. https://doi.org/10.1017/S0142716401004027

CHAPTER 6

Prediction in bilingual children

The missing piece of the puzzle

Figen Karaca[1], Susanne Brouwer[1], Sharon Unsworth[1] and Falk Huettig[1,2]
[1]Radboud University / [2]Max Planck Institute for Psycholinguistics

A wealth of studies has shown that more proficient monolingual speakers are better at predicting upcoming information during language comprehension. Similarly, prediction skills of adult second language (L2) speakers in their L2 have also been argued to be modulated by their L2 proficiency. How exactly language proficiency and prediction are linked, however, is yet to be systematically investigated. One group of language users which has the potential to provide invaluable insights into this link is bilingual children. In this paper, we compare bilingual children's prediction skills with those of monolingual children and adult L2 speakers, and show how investigating bilingual children's prediction skills may contribute to our understanding of how predictive processing works.

What is prediction?

One of the most fascinating characteristics of language comprehension is how efficient and effortless it is in spite of the fast and incremental nature of spoken language. Listeners actively process the rapid speech signals as they unfold by not only incrementally analyzing incoming input but also by generating predictions about the upcoming information. In other words, they successfully pre-activate specific linguistic input before it is encountered (e.g., Altmann & Mirkovic, 2009; Dell & Chang, 2014; Federmeier, 2007; Ferreira & Chantavarin, 2018; Gibson et al., 2013; Hale, 2001; Hickok, 2012; Huettig, 2015; Kuperberg & Jaeger, 2016; Levy, 2008; Norris et al., 2016; Pickering & Gambi, 2018; Pickering & Garrod, 2013; van Petten & Luka, 2012). It is worth noting that there is no consensus on the definition of *prediction* in language. Some argue for differentiating between facilitation and prediction in that the former means faster and easier processing of a word, whereas the latter requires pre-activation of the linguistic representation of a specific word. In this paper, we will include studies that differentiate facilitation from prediction as well as those that do not. We believe that in almost all cases facilitation of a

https://doi.org/10.1075/bpa.12.06kar

word is a consequence of pre-activation (e.g., through priming) and thus part of predictive processing. We define prediction here as the pre-activation of linguistic representations before incoming bottom-up input has had a chance to activate them (Huettig, 2015).

Information from various levels of representation including but not limited to morphosyntax, semantics, and discourse might serve as a reliable cue in predicting the meaning of the upcoming signal. For instance, Altmann and Kamide (1999), in their seminal work, argued that monolingual adult speakers use the semantic restrictions of verbs to predict upcoming information. Using a visual world paradigm (VWP; Cooper, 1974; Tanenhaus et al., 1995), they presented participants with sentences that contained a semantically restraining verb such as *The boy will eat the cake*, or a neutral verb such as *The boy will move the cake* in a visual context of a toy train set, a toy car, a balloon, and a birthday cake. In this context, only the cake was edible while all objects could be moved by the agent. The analyses of the eye-movements revealed that having heard the verb *eat*, the participants looked at the only edible object more often before encountering the word *cake*. Based on these anticipatory eye movements, it was clear that the listeners were able to make predictions about the upcoming information in a sentence based on the cues at their disposal, in this case verb semantics.

How are prediction and proficiency related?

Studies conducted so far have robustly suggested that not only monolingual adults but also monolingual children successfully engage in predictive language processing (e.g., Brouwer, Sprenger, & Unsworth, 2017; Brouwer et al., 2019; Havron et al., 2019; Lew-Williams & Fernald, 2007; Mani & Huettig, 2012; Özge et al., 2019). Prediction skills of monolingual adults and children have been argued to be modulated by their language proficiency, variously measured using target-like production of certain linguistic structures, vocabulary knowledge, and reading skills (e.g., Borovsky et al., 2012; Brouwer, Sprenger, & Unsworth, 2017; Gambi et al., 2020; Huettig & Brouwer, 2015; Mani et al., 2016; Mani & Huettig, 2012, 2014; Rommers et al., 2015). This link between language proficiency and prediction may be bidirectional, in that not only people improve in prediction as they become more proficient language users, but also prediction ability may support linguistic development through facilitating processing of linguistic input in childhood (e.g., Gambi et al., 2020; Gambi, this volume). In other words, prediction may directly or indirectly facilitate language learning (though it may not be necessary for learning, see Huettig & Mani, 2016; see also Hopp, this volume).

The importance placed on prediction in relation to cognition and language learning has led to extensive research on second language (L2) users' engagement in prediction. Given that L2 acquisition is characterized by considerable individual variation, it is not surprising that studies in L2 predictive processing have yielded mixed results. Some studies provided evidence for successful prediction effects to a similar extent as monolingual speakers (e.g., Dijkgraaf et al., 2017; Ito et al., 2018), whereas other studies demonstrated smaller or delayed effects and in some cases no effect of prediction (e.g., Hopp, 2015; Mitsugi & MacWhinney, 2016). The studies reporting little to no prediction effects mostly investigated predictive use of (morpho)syntactic cues such as case and gender marking, which is difficult to master even for monolingual speakers, depending on the transparency of the cues. Adult L2 speakers' ability to use such cues predictively has been shown to be modulated by their L2 proficiency (e.g., Dussias et al., 2013; Hopp, 2013; Hopp & Lemmerth, 2018; but cf. Hopp, 2015; Ito et al., 2018; Dijkgraaf et al., 2017) and the presence of the same cues in their L1 (e.g., Dussias et al., 2013; Foucart et al., 2014; see also Foucart, this volume).

Overall, then, increased language proficiency as measured by vocabulary size, reading skills, and target-like production of certain structures, has been shown to facilitate prediction skills of not only monolingual speakers but also adult L2 speakers. In the case of the latter group, L1-L2 similarity also plays a role.

Not only linguistic but also cognitive skills play a role in predictive processing. This is because language processing and more general cognitive processing are closely intertwined. Studies suggest, for instance, that age-related cognitive changes result in decreased prediction in language processing (e.g., Federmeier & Kutas, 2005), and higher working memory capacity in increased prediction in language processing (Huettig & Janse, 2016; Ito et al., 2018). This suggests that language proficiency and the availability of more general cognitive resources both determine prediction in language processing and one should not be considered without the other.

What we know so far about predictive processing in relation to language proficiency and cognitive skills has been shaped around data provided by monolingual speakers (e.g., Huettig & Janse, 2016; Mani & Huettig, 2012, 2014) and by adult L2 speakers (e.g., Dussias et al., 2013; Hopp, 2013; Hopp & Lemmerth, 2018). One group that has been neglected so far but may contribute to our understanding of prediction in general is bilingual children. Within the framework of this chapter, we define bilingual children as children who were exposed to two languages before the age of four (e.g., Genesee et al., 2004; McLaughlin, 1978; Unsworth, 2013a), and who are exposed to and use both their languages in their daily lives, that is, what are commonly referred to as simultaneous or early sequential bilinguals.

Investigating bilingual children's prediction skills may help us better understand the way predictive processing works. More specifically, this line of investigation may be informative in two different ways: due to both similarities to and differences from (a) monolingual children and (b) adult L2 speakers. First, on average, bilingual children are similar to monolingual children in their trajectory of developing cognitive skills but different in that they are acquiring more than one language. Being exposed to two languages in their daily lives causes significant variation in the language environments of bilingual children both between bilingual children and as a group in comparison to monolingual children. As a result, even though most bilingual children's language development is within normal range in one of their languages and sometimes in both, their language proficiencies show significant variation. The extent of this variation is what sets bilingual children apart from their monolingual peers, who also vary but likely to a lesser extent, and it may enable researchers to investigate the role of language proficiency in predictive processing more comprehensively.

The second way in which investigating bilingual children's prediction skills may inform our understanding of predictive processing is in comparison with adult L2 speakers. Bilingual children are similar to adult L2 speakers in that they have two languages but they differ as to whether these languages develop sequentially or simultaneously. Adult L2 speakers have an already entrenched L1 system when they start to learn an L2. Furthermore, they often learn their L2 in classroom environments, which limits their experiences with the target language significantly. Because they are more dominant and proficient in their L1, they often exhibit unidirectional cross-linguistic influence, that is from L1 to L2. In contrast, bilingual children's two languages develop more or less in parallel and their relative proficiency in the two languages varies considerably: they may be more or less equally proficient in both languages or more proficient in either of their languages. Such varied relative proficiencies and potential effects of bidirectional cross-linguistic influence offer a good place to start exploring the mediating role of language proficiency in prediction skills in one language, and the interaction of prediction skills in two languages in a developing mind.

In sum, these differences show us that including bilingual children into prediction research might provide new insights into prediction that neither monolingual children nor adult L2 speakers offer. This paper will argue that given the similarities and differences between bilingual and monolingual children as well as bilingual children and adult L2 speakers, studies with bilingual children have the potential to provide a relevant test case to investigate (1) the role of language proficiency in prediction due to the significant amount of variation that is observed in bilingual children in comparison to monolingual children and adult L2 speakers, and (2) the role of cross-linguistic influence in predictive processing in the absence of a fully-acquired L1. The following sections will review the studies with monolingual children and adult L2 speakers and show how examining bilingual children's prediction skills could contribute to predictive processing research. We will also show how studies with bilingual children may inform (L2) predictive processing accounts, concluding with suggestions for future research.

What do we know about prediction skills in monolingual children?

Children's predictive language processing skills have predominantly been investigated in monolingual populations. A growing number of studies suggests that monolingual children can exploit cues from various sources such as verb semantics (e.g., Borovsky et al., 2012; Brouwer et al., 2019; Gambi et al., 2018; Mani & Huettig, 2012; Mani et al., 2016) and morphosyntax (e.g., Brouwer, Sprenger, & Unsworth, 2017; Lew-Williams & Fernald, 2007; Lukyanenko & Fisher, 2016; Melançon & Shi, 2015; van Heugten & Shi, 2009) in order to generate predictions about the upcoming information. However, the way in which the prediction system of monolingual children develops is to a certain extent language-specific, in that not all languages have the same predictive cues. Their performance, moreover, is modulated by their language proficiency as measured with vocabulary knowledge and target-like production of certain structures.

Previous work has shown that children may rely on different cues in predictive processing depending on the language they speak. For some languages monolingual children can use a cue predictively at a very young age, whereas in some other languages predictive use of the same cue may be delayed. For instance, monolingual German-speaking children as young as 2-year-old can predict upcoming information when presented with sentences containing semantically restricting or neutral verbs (e.g., *eat* vs. *see*) accompanied by two familiar images (e.g., cake and bird) (Mani & Huettig, 2012). Similarly, monolingual English-speaking children were also able to combine such semantic cues with structural relations (e.g., argument structure) to guide their predictions around the age of four (Gambi et al., 2016). In contrast, 4-to-5-year-old monolingual Turkish-speaking children were not able to use verb semantics predictively unlike their Dutch-speaking peers (Brouwer et al., 2019). Furthermore, when the word order and verb semantics were the only cues at their disposal (i.e., in the absence of any case-marking cues), Turkish monolingual children showed uncertainty in figuring out the argument structure around the ages of one to three (Candan et al., 2012). One possibility explaining these results is that children speaking a head-final language might prioritize early-arriving cues (e.g., case-marking) over late-arriving cues (e.g., verb semantics) (Choi & Trueswell, 2010), or that in such languages morphosyntactic cues bear more predictive power.

In line with such suggestions, it has been observed that Turkish-speaking monolingual children can successfully exploit case-marking cues for predictive processing. Özge and colleagues (2019) investigated whether 4-to-5-year-old Turkish-speaking monolingual children can use accusative or nominative case-marking on the first noun phrase (NP) to predict the second NP in the sentence, where the former marks the direct object while the latter the subject. They presented children with sentences in which the first NP was either in nominative (i.e., the subject) or accusative (i.e., the direct object) case. Accompanying such sentences, a visual context with three related images was presented. These images represented the first NP (e.g., rabbit), a plausible patient in a context where the first NP is the agent (e.g., carrot), and a plausible agent in a context where the first NP is the patient (e.g., fox). The results showed that in verb-final sentences, after hearing the sentence initial accusative-marked NP, children fixated more on the plausible agent prior to hearing the verb and the second NP. These findings clearly demonstrated that children were sensitive to case-marking information, and further used that cue to predict the thematic role of the upcoming noun. Taken together, it appears to be the case that the strategies monolingual children employ in predictive processing depend heavily on the language they speak.

In addition to verb semantics and case-marking, monolingual children have also been found to exploit gender-marking cues predictively. In one of the key studies, Lew-Williams and Fernald (2007) examined whether 3-year-old monolingual Spanish-speaking children were able to benefit from gender-marking on articles in predicting the upcoming noun, using a looking-while-listening task. The results revealed that children identified the target image faster when the gender of the target and the distractor image were different, thus informative about the referent of the upcoming noun. These findings suggested that gender cues facilitated language processing in Spanish-speaking monolingual children. This facilitation effect showed that children were able to make predictions about the upcoming nouns based on gender cues on the preceding articles in Spanish, which has a transparent gender-marking system. In contrast, 2-year-old Dutch-speaking monolingual children were found to experience difficulties with processing such cues, in that they could use the common gender-marked article *de* predictively, but not the neuter gender-marked *het* (Johnson, 2005). As they get older and become more adult-like in production of gender-marking in Dutch, they do however start to use both gender cues predictively (Brouwer, Sprenger, & Unsworth, 2017; cf. Kochari & Flecken, 2019). Observing the differences between the Spanish- and Dutch-speaking children, it is conceivable that the nature of the gender-marking system in different languages (transparent versus opaque) affects whether and, if so, how such information can be processed predictively in different languages (though further research is needed to replicate these differences between Spanish- and Dutch-speaking children).

In sum, the available research demonstrates that monolingual children can predict upcoming information in a sentence using a diverse number of cues including verb semantics, case- and gender-marking. The way in which their prediction skills develop and their level of sensitivity towards different predictive cues appear to depend (at least partly) on the characteristics of the language in question.

Not only language-level but also individual-level factors such as the proficiency of the monolingual children play a role in the development of prediction skills. Language proficiency may be indexed by receptive and/or productive vocabulary size and target-like production of certain linguistic structures, and it is closely related to the children's language environment. Accumulating empirical evidence strongly suggests an association between vocabulary knowledge and predictive processing of verb semantics in monolingual children (Borovsky et al., 2012; Gambi et al., 2020; Mani & Huettig, 2012; Mani et al., 2016; cf. Gambi et al., 2016). For example, Borovsky and colleagues (2012) reported that 3-to-10-year-old monolingual children and adults with higher receptive vocabulary skills were faster in prediction. Furthermore, Mani and Huettig (2012) argued that the number of words that 2-year-old monolingual children were able to not only understand but also produce correlated positively with their prediction skills. Similar effects of production abilities were also attested in Brouwer, Sprenger, and Unsworth (2017), where monolingual Dutch-speaking children with target-like production of gender agreement were able to use gender cues predictively, whereas those with non-target-like production were only able to use the same cues facilitatively. That is, target-like production might have triggered successful online comprehension in the form of prediction. Taken together, the available research suggests that the language proficiency of monolingual children, as measured by their vocabulary knowledge and target-like production performance, is associated with their prediction skills.

In line with these studies, Mani et al. (2016) suggested that it is not only language proficiency but also language experience which affects monolingual children's prediction skills. More specifically, these authors found that children with a larger productive vocabulary were better in predicting the words that were strongly (e.g., *book*) or weakly associated (e.g., *letter*) with the verb (e.g., *read*) compared to an unassociated word (e.g., *cheese*). However, a significant correlation between prediction skills and productive vocabulary disappeared in trials in which weakly and strongly associated objects were presented together in the visual context (e.g., *the boy reads the book* vs. *the letter*). In such cases, prediction performance was modulated by the relative associative strength of the words in relation to a specific verb based on past experiences. In other words, the prediction skills of monolingual children appear to be influenced by their language experience (Foucart, 2015; Mani & Huettig, 2014). The role of language experience may be more apparent in languages with relatively more opaque systems (e.g., gender marking in Dutch) since mastery and predictive use of opaque structures require substantial linguistic input (Unsworth, 2013b).

Overall, it has been suggested that larger receptive and productive vocabulary size and target-like production abilities (i.e., higher language proficiency) may make monolingual children better predictors. More and diverse language experience of monolingual children benefits their proficiency in language, facilitating their prediction skills.

The variation observed among monolingual children highlights the complexity of prediction and significance of language proficiency, necessitating more comprehensive research into how exactly proficiency modulates prediction skills. However, there is only so much variation among monolingual children in terms of language proficiency and experience that will allow researchers to properly investigate the role of such factors in prediction. Bilingual children, however, are likely to show comparatively more variation in their linguistic skills and language environments, and therefore offer an interesting test case. Assuming that all typically developing children have similar cognitive resources available during language comprehension, and that bilingual children (as a group) show more variation in their language experience and proficiency when compared to monolingual children, bilingual children's prediction skills will offer invaluable insights into the exact role of these factors in predictive processing.

Not only differences in proficiency but also language-specific properties of bilingual children's other language may affect predictive processing. The next section reveals that, based on what we know about adult L2 speakers, there are reasons to believe this might be the case.

What do we know about prediction skills in adult L2 speakers?

The research on adult L2 speakers' prediction skills so far has yielded mixed findings. Some studies provided evidence for successful prediction effects to a similar extent as monolingual speakers (e.g., Dijkgraaf et al., 2017), whereas other studies demonstrated smaller or delayed and in some cases no effect of prediction (e.g., Hopp, 2015; Mitsugi & MacWhinney, 2016). There are a multitude of interrelated factors that modulate L2 predictive processing of which cross-linguistic influence and L2 proficiency have been shown to be of major impact.

One of the factors involved in adult L2 prediction abilities is the interaction between the two languages of bilinguals. Due to the non-selective nature of bilingual language processing, linguistic input from one language co-activates both languages. In the course of language processing, two concomitantly active languages interact with each other, not only at the lexical but also at the syntactic level (e.g., Kootstra & Doedens, 2016). For instance, studies have shown that bilingual speakers recognize and produce cognate words faster than non-cognates (see Lijewska, 2020, for a detailed overview), and that their choice of syntactic structures and processing strategies are under the influence of the other language (see van Gompel & Arai, 2018). Co-activation of languages might slow down or facilitate predictive processing, as successful prediction requires attention and sensitivity towards a variety of language-specific cues.

For instance, Martin and colleagues (2013) investigated whether highly proficient Spanish-English adult L2 speakers engaged in prediction while reading highly constrained sentences in which the final noun (expected versus unexpected and starting with a vowel versus consonant) and the preceding article (*a* versus *an*) were manipulated (e.g., *Since it is raining, it is better to go out with an umbrella/ a raincoat*). They employed an ERP paradigm in which the lexical prediction effect was indexed by the N400 effect elicited by the unexpected article. The results revealed a greater N400 effect only in the monolingual group. The lack of such an effect in L2 speakers was interpreted as a failure to predict the target noun. One important confound that was overlooked in relation to these results was the difference in the article systems of the two languages. Even though both Spanish and English have articles, their selection in Spanish is not driven by phonological properties of the following noun. It is, then, possible that the L2 speakers might still be able to predict the following noun, but failed to show the prediction effect on the article in English. It is also important to note that the prediction effect that was found in the monolingual group in Martin et al. (2013) was not replicated in later studies (e.g., Nieuwland et al., 2018).

Using a similar paradigm, Foucart and colleagues (2014) reported N400 effects elicited by the article that matched the gender of the unexpected noun in both monolinguals as well as French-Spanish adult L2 speakers. The authors argued that their contradictory findings resulted from cross-linguistic similarities and differences as the L2 speakers in their study were able to use a morphosyntactic cue (i.e., gender-marking on the article) that was readily available in both languages. Similar facilitative effects of cross-linguistic similarity in the gender system were also reported in Dussias et al. (2013) with regard to Italian-Spanish adult bilinguals. It should however be noted that all the critical nouns in Foucart et al. (2014) were carefully selected so that they bore the same gender in both languages. Therefore, it is hard to identify the exact source of the observed prediction effect: Is it the presence of a gender system in both languages, or the gender overlap between lexical items and their translation equivalents?

The role of gender overlap was investigated in a recent study with Russian-German adult L2 speakers (Hopp & Lemmerth, 2018). Both Russian and German are gender-marking languages. Importantly, the nouns in both languages may bear the same or different genders, and both languages mark gender agreement on adjectives, whereas only German does so on articles. The findings of this study showed that adult L2 speakers of German with high-intermediate level of proficiency were able to use gender-marking on adjectives predictively regardless of whether the target noun bore the same gender in both languages. However, they showed predictive use of gender-marking on articles only when genders of the target noun overlapped in the two languages. These findings suggested that gender overlap benefitted adult L2 speakers specifically when gender was marked syntactically different in L1 and L2.

In addition to gender-marking cues, the predictive use of another morphosyntactic cue, namely case-marking, has also been examined in adult L2 speakers. Using a VWP similar to Kamide et al. (2003) where case-marking cues and verb semantics carried predictive information about the second NP, Hopp (2015) tested monolingual German speakers and English-German adult L2 speakers. He found that monolingual speakers fixated on the target image before the onset of the second NP in both SVO and OVS sentences, suggesting that they were able to integrate information from case-marking and verb semantics. In contrast, adult L2 speakers fixated on the patient image regardless of the case-marking on the first NP. This finding suggested that L2 speakers were not able to employ case-marking cues predictively, instead they relied only on verb semantics. Corroborating these findings, English-Japanese L2 speakers were also shown to be unable to generate predictions based on case-marking cues (Mitsugi & MacWhinney, 2016).

Compared to morphosyntactic cues, verb semantics appears to be less susceptible to cross-linguistic differences, most likely due to its reliance on general world-knowledge. For example, Dijkgraaf and colleagues (2017, 2019) reported successful prediction effects in Dutch-English L2 speakers. Some aspects of verb semantics, however, are more language-specific. For instance, German and Dutch encode specific positional information in placement verbs (i.e., *zetten* for standing objects, and *leggen* for lying objects in Dutch). It has been suggested that this language-specific information was used predictively by monolingual Dutch-speaking adults and by German-Dutch L2 speakers. In contrast, L2 speakers whose L1 did not specify such information such as French and English, were found to not use the same cue predictively (van Bergen & Flecken, 2017). The similarity between Dutch and German in encoding the position of the object in placement verbs was argued to facilitate predictive processing. These findings indicate that not even the cues from verb semantics are immune to cross-linguistic influence.

In sum, cross-linguistic similarities and differences play a key role for adult L2 speakers when it comes to using cues from morphosyntax or verb semantics predictively. Due to the co-activation of both languages, predictive processing in the L2 might be affected by the presence of conflicting cues, or it may be facilitated by the similarities in the processing strategies.

Similar to monolingual children, individual-level factors such as language proficiency, language experience and cognitive skills, also play a significant role in predictive processing in adult L2 speakers. One obvious and important determinant of L2 predictive processing is L2 proficiency (e.g., Dussias et al., 2013; Hopp & Lemmerth, 2018; cf. Hopp, 2015). For instance, Hopp and Lemmerth (2018) found that unlike high-intermediate Russian-German adult L2 speakers, advanced L2 speakers were able to use gender cues predictively regardless of gender overlap or differences in the syntactic realization of gender between the two languages. Moreover, Hopp (2013) found that English-German adult L2 speakers who assigned correct gender on nouns in a production task were able to use gender cues predictively, whereas the ones with less target-like gender assignment were not able to do so. These findings underscore that L2 proficiency significantly modulates predictive processing. Relatedly, Foucart (2015) pointed out that increased L2 experience may help adult L2 speakers, such that increased familiarity with L2 structures and co-occurrences may benefit predictive processing. Lastly, predictive processing in L2 was argued to depend on the availability of cognitive resources and skills of listeners (Ito et al., 2018).

In short, then, it is evident that prediction skills of L2 speakers show substantial variation and the ability to generate predictions is modulated by not only language-level (i.e., cross-linguistic influence) but also individual-level factors (e.g., L2 proficiency, language experience and cognitive skills). However, our knowledge and assumptions about how prediction occurs in one language in the presence of another language are based almost exclusively on data provided by adult L2 speakers. Even though adult L2 speakers have the advantage of cognitive maturity, they have acquired their L2 with an entrenched L1 system and often have relatively limited L2 experience. They are typically dominant and more proficient in their L1, which makes unidirectional cross-linguistic influence likely when it comes to predictive processing. In contrast, bilingual children acquire two languages in parallel. Their language experiences and relative proficiencies in two languages vary significantly, making bidirectional cross-linguistic influence also a more plausible option in their case. Because bilingual children's relative language proficiency varies significantly, they are more likely to spread across the full 'continuum of bilingualism' (Luk & Bialystok, 2013). More widely distributed positions of bilingual children on this continuum may enable researchers to examine the effects of proficiency in prediction in more detail. Therefore, investigating bilingual children's prediction skills might offer a new perspective in prediction research in terms of the role of language proficiency and cross-linguistic influence.

What do we (not) know about bilingual children?

Despite the wealth of studies with monolingual children and adult L2 speakers, research on bilingual children's prediction skills is newly emerging and rather scarce (Brouwer, Özkan, & Küntay, 2017; Lemmerth & Hopp, 2019; Meir et al., 2020). Nevertheless, the research that has been done has yielded interesting and promising findings.

In one of the first prediction studies with bilingual children, Brouwer, Özkan and Küntay (2017) investigated whether 4–5-year-old bilingual children with heterogeneous L1 backgrounds could use verb semantics cues predictively compared to monolingual Dutch-speaking children. In a VWP similar to the one used in Mani and Huettig (2012), they presented children with sentences containing a semantically constraining (e.g., *eat*) or neutral (e.g., *see*) verb, accompanied by two pictures (e.g., cake and tree). The results showed that bilingual children were able to make predictions about upcoming information on the basis of verb semantics in Dutch similar to their monolingual peers. In fact, the 4-year-old bilingual children outperformed their monolingual peers. These findings suggested that bilingual children were in principle able to use verb-semantics predictively to the same extent as their monolingual peers, if not better.

Two more recent studies have examined bilingual children's prediction skills using morphosyntactic cues. Lemmerth and Hopp (2019), for example, investigated whether 7-to-9-year-old Russian-German simultaneous bilingual children and bilingual children with an age of onset between one to three years were able to use gender-marking cues predictively, by adapting Hopp and Lemmerth's (2018) VWP experiment for children. The analyses of the reaction times revealed no qualitative differences between prediction skills of monolingual and simultaneous bilingual children, though the latter group was slower overall. However, bilingual children who were exposed to German between the ages of one to three launched earlier looks to the target picture only when the target noun bore the same gender in the two languages. These results demonstrated a predictive processing effect that was modulated by the lexical gender-congruency between languages. This effect indicated that linguistic input co-activated both of the languages these bilingual children were acquiring. The authors argued that hearing a gender-marked article in one language activated all nouns bearing the same gender across languages. As a result, the nouns sharing the same gender with their translation equivalents benefitted from this non-selective co-activation and were thus more easily predicted. This study therefore suggested that lexical gender overlap may aid predictive use of gender cues for bilingual children who were exposed to German between the ages of one to three, whereas the gender-incongruent nouns may suffer from competition effects.

The second study investigating predictive processing in bilingual children focussed on case-marking cues. Meir and colleagues (2020) examined whether 4-to-8-year-old Russian-Hebrew bilingual children were able to employ case-marking cues in predictive processing, in comparison to monolingual Russian- and Hebrew-speaking children. In a VWP, monolingual Russian-speaking children looked at the plausible agent (e.g., fox) as soon as they heard the accusative-marked NP (e.g., bunny), whereas their Hebrew-speaking peers failed to do so. Bilingual children also used accusative case on the first NP to predict the upcoming agent in Russian, though more slowly than their monolingual peers. Interestingly, they also employed the same cue predictively in Hebrew, after hearing the verb. These results showed that bilingual Russian-Hebrew-speaking children were able to exploit a cue that cannot be used predictively by their monolingual Hebrew-speaking peers. These findings suggested that interaction between case-marking cues from both languages affected predictive processing in bilingual children. The predictive power of a relatively less reliable cue in one language (i.e., Hebrew) was boosted by the presence of the same, but stronger, cue in the other language (i.e., Russian), which is in line with offline studies showing cross-linguistic influence in the form of acceleration relative to monolingual peers (e.g., Meroni et al., 2017).

Overall, the limited number of studies available on bilingual children's prediction skills show mixed findings. In some cases, bilingual children outperformed their monolingual peers (Brouwer, Özkan, & Küntay, 2017; Meir et al., 2020), whereas in other cases their performance aligned with or showed a different pattern from monolingual children (Lemmerth & Hopp, 2019). Such findings highlight that there is so much variation to explore in the case of bilingual children, and that their developing skills may change our assumptions about predictive processing. For instance, knowing another language may not necessarily impede prediction skills, not even in the use of case-marking cues predictively. In short, the available findings of the limited studies with bilingual children show that there is still so much that we do not know, and what we do know is not completely clear, and yet bilingual children could offer a new perspective into predictive processing.

How can research with bilingual children inform L2 predictive processing accounts?

Based on the results of earlier studies reporting limited effects of prediction even for highly proficient adult L2 speakers, Grüter and colleagues proposed that L2 speakers have a reduced ability to generate expectations about the upcoming information because they exhaust almost all of their processing resources on integrating the incoming information (Grüter et al., 2014; Grüter et al., 2017; for a similar explanation also see Pickering & Gambi, 2018). In more recent work, Grüter et al. (2018) refined this hypothesis by stating that the differences between L1 and L2 processing were more likely to be gradual, rather than categorical. L2 speakers may weigh cues (semantic versus form-class) differently than L1 users. In contrast, Kaan (2014) assumes no qualitative differences between L1 and L2 predictive processing mechanisms, but rather highlights the potential role of mediating factors which are yet to be systematically investigated, even in monolingual populations (Huettig, 2015). More specifically, she argues that the mechanisms involved in L1 and L2 predictive processing are fundamentally the same, and that similar mediating factors and individual differences could be responsible for the observed differences between L1 and L2 speakers as well as among L1 speakers. These mediating factors include frequency information about the likelihood of occurrences of words and of structures in a specific context, the quality of lexical representations, as well as motivation, emotional state and cognitive resources. Whilst there are clear differences between the RAGE Hypothesis and Kaan's (2014) account in terms of the underlying reasons behind the observed differences between the two groups, the two accounts essentially make the same prediction when it comes to L2 speakers, namely that more proficient L2 speakers will be more likely to engage in predictive processing.

These two accounts, as well as the available evidence from monolingual children and adult L2 speakers, suggest that language proficiency may modulate predictive processing in L1 and L2. We argue that the exact role of language proficiency may be better understood when investigated in bilingual children. We hypothesize that, as for monolingual children and adult L2 speakers, the language proficiency of bilingual children influences their prediction skills in the language in question. Less proficient listeners might need to allocate more cognitive resources during language processing, leaving limited resources to generate predictions. That means that depending on their proficiency in that language, bilingual children may or may not engage in predictive processing to the same extent as their monolingual peers. Not only differences in proficiency but also language-specific properties of bilingual children's other language may affect predictive processing. We furthermore hypothesize that cross-linguistic influence might take place at the level of prediction since bilingual children's two languages could employ similar as well as different predictive strategies. In the remainder of this paper, we will discuss how exactly research into bilingual children's prediction skills may help us better understand the relationship between prediction and proficiency.

Even though language acquisition proceeds in similar ways in monolingual and bilingual children, there are important differences between the language environments of monolingual and bilingual children. More specifically, there are factors influencing the quantity and quality of language exposure children receive which may differ across bilingual and monolingual contexts. To varying degrees, these factors modulate how proficient bilingual children are in two languages. This means that if proficiency is central to (developing) prediction skills, then the quantity and quality of exposure in bilingual children's two languages should also be related to their prediction skills.

Input quantity has been previously argued to affect bilingual children's language outcomes (see Paradis, 2011; Unsworth, 2016, for an overview) as well as their processing skills (Sorace, 2005). For instance, cumulative length of exposure has been reported to modulate the production of gender agreement in Dutch (Unsworth, 2013b), and the comprehension of wh-questions disambiguated by case-marking cues in German, depending on the position and number of such cues available in the sentence (Roesch & Chondrogianni, 2016). In a study on bilingual Spanish-English toddlers, Hurtado et al. (2014) observed a complex relationship between amount of exposure, processing speed and vocabulary size. They argued that children who had comparatively larger vocabulary sizes in one language were able to process words in that language more quickly (as measured by mean RTs in one language divided by the mean RTs in the other language), and the relative speed of processing words in one language was tied to relative experience in that language. In other words, relatively more exposure in one language increases bilingual children's experience and practice in that language, which promotes their language processing skills. According to the authors, children with increased processing speed may take better advantage of the linguistic input that they receive, which subsequently helps them learn new vocabulary faster in that language. In turn, the larger vocabulary size in one language makes language processing easier and faster for children in that language. Given the available evidence relating language exposure to language proficiency and processing, we expect that amount of exposure will also be positively related to bilingual children's predictive processing skills.

In addition to quantity of input, the quality of input may also play a role in bilingual children's prediction skills. The quality of input encompasses various factors including the number of input providers in the language environment and their nativelikeness, as well as the richness of input (e.g., language activities such as reading) (Paradis, 2011). For instance, Hoff et al. (2020) found that lexical and grammatical features of the child-directed speech provided by native speakers and proficient non-native speakers significantly differed from that of non-native speakers with limited proficiency. In other words, proficiency of non-native speakers modulated the richness of the input that children received. Relatedly, Unsworth et al. (2019) found that receptive vocabulary and morphosyntactic skills of 3-year-old bilingual children with heterogenous L1 backgrounds were modulated by the degree of nativelikeness of the input providers. The authors argued that morphosyntactically complex and lexically diverse input provided by more proficient non-native speakers helped bilingual children's language outcomes. With respect to prediction skills, we can derive the following hypotheses from these findings. The more bilingual children are exposed to consistent and rich linguistic input due to the high proficiency level of input providers, the more likely that they will be able to notice and derive correct structures from the input. Such input is more likely to increase the strength of associations between words and structures for bilingual children, making possible predictive cues more reliable.

Another measure of the input quality is richness of the language environment, which is often indexed by language activities such as reading (Paradis, 2011). Enriching bilingual children's language experience through reading activities may benefit their language outcomes as well as their prediction skills. Such an association between reading skills and prediction has been previously suggested (see Huettig & Pickering, 2019) in relation to prediction skills of monolingual children (Mani & Huettig, 2014) and adults (Kukona et al., 2016), monolingual adults with dyslexia (Huettig & Brouwer, 2015), and low and high literates (Mishra et al., 2012) partly because reading activities are argued to enhance the quality of the linguistic representations, making prediction during language comprehension more viable.

In addition to leading to a better understanding of the relation between proficiency and prediction, bilingual children can also help us further understand the interaction between cross-linguistic influence and language proficiency in prediction. The parallel acquisition of two languages, coupled with varied relative proficiencies in two languages may result in differences in prediction skills of bilingual children compared to adult L2 speakers, which may be informative on how language proficiency interacts with cross-linguistic influence in predictive processing. For instance, target-like production ability, which is an indicator of language proficiency, has been found to play a role in predictive processing skills of both monolingual children (Brouwer, Sprenger, & Unsworth, 2017) and adult L2 speakers (Hopp, 2013); it seems logical, then, to expect similar effects in bilingual children. Nonetheless, Meir et al. (2020) found that Russian-Hebrew bilingual children, who were less accurate in production of accusative case morphology both in Russian and in Hebrew, were able to use accusative case-marking on the first NP to predict the upcoming second NP in online language processing. The authors argued that the prediction skills of these bilingual children in Hebrew benefitted from the presence of stronger case-marking cues in their other language (i.e., Russian). The effects of non-target like production in this case may have been mitigated by the strong interaction of the same predictive cues in two languages during language processing. In other words, cross-linguistic influence in the form of acceleration took place in predictive language processing. It is important to note here that even though the topic of cross-linguistic influence in bilingual children is well researched (see van Dijk et al., in press, for meta-analysis), there is very little research on cross-linguistic influence in real-time language processing (but see van Dijk et al., under review).

Since the bilingual children in Meir et al. (2020) were able to use case-marking cues predictively, their less target-like performance in production of case-marking suggests a production-specific problem. This finding is of relevance for predictive processing accounts which specifically argue for the involvement of the production system in prediction. For instance, Pickering and Gambi (2018) have argued that using production systems is the most effective way of predicting; however, because this route takes time and resources, non-native speakers may use it less, which makes it an optional mechanism for less proficient language users (see also Ito & Pickering, this volume). The findings that showed that bilingual children, who demonstrated a production-specific problem, were able to use case-marking cues in prediction while their monolingual peers were not, may be interpreted to suggest that prediction-by-production is not always the most effective way of predicting.

What we know about bilingual children's prediction skills so far shows that knowing another language is not necessarily a disadvantage in predictive processing. The amount and quality of input that bilingual children receive in each language varies significantly, which in turn is an important factor predicting their relative proficiency in their two languages. The considerable variation observed in bilingual children's relative language proficiency spreads them out more widely on the full continuum of language proficiency. Since they are likely to inhabit more diverse positions on this continuum with their varying relative proficiencies compared to adult L2 speakers, investigating their prediction skills may help us understand the relation between proficiency and prediction more comprehensively.

What's next?

Bilingual children offer an interesting case to investigate how language proficiency, experience and cognitive skills as well as cross-linguistic influence affect prediction skills, due to their distinct characteristics in comparison to monolingual children and adult L2 speakers, the two groups which have thus far dominated the predictive processing research. Bilingual children's prediction skills should be systematically investigated by taking individual-level and language-level differences into consideration.

In terms of individual-level differences, it is essential to test prediction skills of bilingual children in both languages and to measure their proficiencies in both languages because relative language proficiency may significantly modulate predictive processing skills of bilingual children in each language. By adopting a within-subjects design, the research with bilingual children may further unravel how language proficiency and experience modulate prediction skills of bilingual children in each language while keeping the cognitive skills constant. It is also important to include different groups of bilingual speakers (i.e., simultaneous, successive and adult L2 speakers) as the amount and timing of exposure to another language may affect how prediction skills develop significantly.

In terms of language-level differences, comparative studies with several groups of bilinguals with different language pairs (e.g., Turkish-Dutch and German-Dutch) will be informative as to the exact nature of cross-linguistic influence in predictive processing (see also van Dijk et al., in press, for a similar argument). Therefore, future studies should focus on investigating prediction skills of bilingual children with diverse L1-L2 pairings.

Acknowledgments

We would like to thank the editors and two anonymous reviewers for their constructive comments and suggestions. We are also grateful to the 2in1 project members for their helpful comments on earlier versions of this chapter.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583–609. https://doi.org/10.1111/j.1551-6709.2009.01022.x

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, 112(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Brouwer, S., Özkan, D., & Küntay, A. C. (2017). Semantic prediction in monolingual and bilingual children. In E. Blom, L. Cornips, & J. Schaeffer (Eds.), *Cross-linguistic influence in bilingualism*, (pp. 48–73). John Benjamins. https://doi.org/10.1075/sibil.52.04bro

Brouwer, S., Özkan, D., & Küntay, A. C. (2019). Verb-based prediction during language processing: The case of Dutch and Turkish. *Journal of Child Language*, 46(1), 80–97. https://doi.org/10.1017/S0305000918000375
Brouwer, S., Sprenger, S., & Unsworth, S. (2017). Processing grammatical gender in Dutch: Evidence from eye movements. *Journal of Experimental Child Psychology*, 159, 50–65. https://doi.org/10.1016/j.jecp.2017.01.007
Candan, A., Küntay, A. C., Yeh, Y., Cheung, H., Wagner, L., & Naigles, L. R. (2012). Language and age effects in children's processing of word order. *Cognitive Development*, 27(3), 205–221. https://doi.org/10.1016/j.cogdev.2011.12.001
Choi, Y., & Trueswell, J. C. (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, 106(1), 41–61. https://doi.org/10.1016/j.jecp.2010.01.003
Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology*, 6(1), 84–107. https://doi.org/10.1016/0010-0285(74)90005-X
Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1634), 20120394. https://doi.org/10.1098/rstb.2012.0394
Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547
Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2019). Prediction and integration of semantics during L2 and L1 listening. *Language, Cognition and Neuroscience*, 34(7), 881–900. https://doi.org/10.1080/23273798.2019.1591469
Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915
Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x
Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, 42(2), 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x
Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, 27, 443–448. https://doi.org/10.1177/0963721418794491
Foucart, A. (2015). Prediction is a question of experience. *Linguistic Approaches to Bilingualism*, 5(4), 465–469. https://doi.org/10.1075/lab.5.4.04fou
Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1461–1469. https://doi.org/10.1037/a0036756
Gambi, C., Gorrie, F., Pickering, M. J., & Rabagliati, H. (2018). The development of linguistic prediction: Predictions of sound and meaning in 2- to 5-year-olds. *Journal of Experimental Child Psychology*, 173, 351–370. https://doi.org/10.1016/j.jecp.2018.04.012
Gambi, C., Jindal, P., Sharpe, S., Pickering, M. J., & Rabagliati, H. (2020). The relation between preschoolers' vocabulary development and their ability to predict and recognize words. *Child Development*. https://doi.org/10.1111/cdev.13465

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. *Cognition*, 157, 340–351. https://doi.org/10.1016/j.cognition.2016.10.003

Genesee, F., Paradis, J., & Crago, M. (2004). *Dual language development and disorders: A handbook on bilingualism and second language learning*. Baltimore, MD: Brookes.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). The rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, 110(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Grüter, T., Lau, E., & Ling, W. (2018). L2 listeners rely on the semantics of classifiers to predict. In A. B. Bertolini & M. J. Kaplan (Eds.), *Proceedings of the 42nd Annual Boston University Conference on Language Development* (pp. 303–316). Cascadilla Press.

Grüter, T., Rohde, H., & Schafer, A. J. (2014). The role of discourse-level expectations in non-native speakers' referential choices. In W. Orman & M. J. Valleau (Eds.), *Proceedings of the 38th Annual Boston University Conference on Language Development* (pp. 179–191). Cascadilla Press.

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, 7(2), 199–229. https://doi.org/10.1075/lab.15011.gru

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of North American Chapter of the Association for Computational Linguistics on Technologies* (pp. 159–166). Stroudsburg, PA: Association for Computational Linguistics. https://doi.org/10.3115/1073336.1073357

Havron, N., de Carvalho, A., Fiévet, A. C., & Christophe, A. (2019). Three-to four-year-old children rapidly adapt their predictions and use them to learn novel word meanings. *Child Development*, 90(1), 82–90. https://doi.org/10.1111/cdev.13113

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13, 135–145. https://doi.org/10.1038/nrn3158

Hoff, E., Core, C., & Shanks, K. F. (2020). The quality of child-directed speech depends on the speaker's language proficiency. *Journal of Child Language*, 47, 132–145. https://doi.org/10.1017/S030500091900028X

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and morphosyntax in L2 predictive sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014

Hopp, H. & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, 40(1), 171–199. https://doi.org/10.1017/S0272263116000437

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Brouwer, S. (2015). Delayed anticipatory spoken language processing in adults with dyslexia – Evidence from eye-tracking. *Dyslexia*, 21(2), 97–122. https://doi.org/10.1002/dys.1497

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huettig, F., & Mani, N. (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, 31(1), 19–31. https://doi.org/10.1080/23273798.2015.1072223

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Science*, 23(6), 464–475. https://doi.org/10.1016/j.tics.2019.03.008

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition*, 17(1), 189–202. https://doi.org/10.1017/S136672891300014X

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Johnson, E. K. (2005). Grammatical gender and early word recognition in Dutch. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development* (pp. 320–30). Cascadilla Press.

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32, 37–55 https://doi.org/10.1023/A:1021933015362

Kochari, A. R. & Flecken, M. (2019). Lexical prediction in language comprehension: a replication study of grammatical gender effects in Dutch. *Language, Cognition and Neuroscience*, 34(2), 239–253. https://doi.org/10.1080/23273798.2018.1524500

Kootstra, G. J., & Doedens, W. J. (2016). How multiple sources of experience influence bilingual syntactic choice: Immediate and cumulative cross-language effects of structural priming, verb bias, and language dominance. *Bilingualism: Language and Cognition*, 19(4), 710–732. https://doi.org/10.1017/S1366728916000420

Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., Pugh, K. R., Shankweiler, D. P., Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica*, 171, 72–84. https://doi.org/10.1016/j.actpsy.2016.09.009

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lemmerth, N. & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition*, 26(1), 21–45. https://doi.org/10.1080/10489223.2017.1391815

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18(3), 193–198. https://doi.org/10.1111/j.1467-9280.2007.01871.x

Lijewska, A. (2020). Cognate processing effects in bilingual lexical access. In R. Heredia & A. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution* (pp. 71–95). Cambridge University Press. https://doi.org/10.1017/9781316535967.005

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two-and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition*, 146, 349–370. https://doi.org/10.1016/j.cognition.2015.10.012

Mani, N., Daum, M. M., & Huettig, F. (2016). "Pro-active" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. *Quarterly Journal of Experimental Psychology*, 69(11), 2189–2201. https://doi.org/10.1080/17470218.2015.1111395

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843–847. https://doi.org/10.1037/a0029284

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology*, 126, 264–279. https://doi.org/10.1016/j.jecp.2014.05.004

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McLaughlin, B. (1978). *Second-language acquisition in childhood*. Lawrence Erlbaum Associates.

Meir, N., Parshina, O., & Sekerina, I. A. (2020). The interaction of morphological cues in bilingual sentence processing: An eye-tracking study. In M. Brown & A. Kohut (Eds.), *Proceedings of the 44th Annual Boston University Conference on Language Development* (pp. 376–389). Cascadilla Press.

Melançon, A., & Shi, R. (2015). Representations of abstract grammatical feature agreement in young children. *Journal of Child Language*, 42(6), 1379–1393. https://doi.org/10.1017/S0305000914000804

Meroni, L., Smeets, L. & Unsworth, S. (2017). Cross-linguistic influence in scope ambiguity: Evidence for acceleration. In E. Blom, L. Cornips, & J. Schaeffer (Eds.), *Cross-linguistic influence in bilingualism: In honor of Aafke Hulk* (pp. 181–206). John Benjamins. https://doi.org/10.1075/sibil.52.09mer

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1). https://doi.org/10.16910/jemr.5.1.3

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. a, Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, 7, e33468. https://doi.org/10.7554/eLife.33468

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience*, 31(1), 4–18. https://doi.org/10.1080/23273798.2015.1081703

Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition*, 183, 152–180. https://doi.org/10.1016/j.cognition.2018.10.026

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213–237. https://doi.org/10.1075/lab.1.3.01par

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Roesch, A. D., & Chondrogianni, V. (2016). "Which mouse kissed the frog?" Effects of age of onset, length of exposure, and knowledge of case marking on the comprehension of wh-questions in German-speaking simultaneous and early sequential bilingual children. *Journal of Child Language*, 43(3), 635–661. https://doi.org/10.1017/S0305000916000015

Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, 77, 720–730. https://doi.org/10.3758/s13414-015-0873-x

Sorace, A. (2005). Syntactic optionality in language development. In L. Cornips & K. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social* (pp. 55–80). John Benjamins. https://doi.org/10.1075/cilt.265.04sor

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, 268(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Unsworth, S. (2013a). Assessing age of onset effects in (early) child L2 acquisition. *Language Acquisition*, 20(2), 74–92. https://doi.org/10.1080/10489223.2013.766739

Unsworth, S. (2013b). Assessing the role of current and CUMULATIVE exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, 16(1), 86–110. https://doi.org/10.1017/S1366728912000284

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan* (pp. 103–122). Mouton de Gruyter. https://doi.org/10.1515/9783110341249-008

Unsworth, S., Brouwer, S., de Bree, E., & Verhagen, J. (2019). Predicting bilingual preschoolers' patterns of language development: Degree of non-native input matters. *Applied Psycholinguistics*, 40(5), 1189–1219. https://doi.org/10.1017/S0142716419000225

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Dijk, C. N., Dijkstra, T., & Unsworth, S. (under review). Cross-linguistic influence during online sentence processing in bilingual children.

Van Dijk, C. N., van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (in press). Cross-linguistic influence in simultaneous bilingual children: A meta-analysis. *Journal of Child Language*. https://doi.org/10.1017/S0305000921000337

Van Gompel, R., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, 21(3), 448–455. https://doi.org/10.1017/S1366728917000542

van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, 12(3), 419–425. https://doi.org/10.1111/j.1467-7687.2008.00788.x

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83(2), 176–190. https://doi.org/10.1016/j.ijpsycho.2011.09.015

CHAPTER 7

Code-switching

A processing burden, or a valuable resource for prediction?

Aleksandra Tomić and Jorge Valdés Kroff
UiT The Arctic University of Norway / University of Florida

Monolinguals use various linguistic phenomena to guide prediction while comprehending. For bilinguals, the richer linguistic landscape provides additional resources. Code-switches (CS) are a particularly salient event, which could play a role in bilingual prediction. Despite the ubiquity and diverse functions of code-switching, experimental research has focused on CS processing costs, largely in comprehension (Litcofsky & Van Hell, 2017). Despite apparent integration costs, code-switching can facilitate subsequent language processing, due to natural code-switching patterns. We illustrate this approach with two eye-tracking studies suggesting that code-switches are used as a cue that a less frequent or negative word follows. These studies underscore the need to integrate socio-pragmatic and corpus-modeling observations with experimentation to reach a comprehensive understanding of CS processing (Myers-Scotton, 2006).

Introduction

Among certain bilingual communities, bilinguals amongst themselves will engage in sentential code-switching, a skillful speech act where bilinguals intentionally switch between languages within the same conversation. Theoretical linguists and sociolinguists have studied the structural and social constraints of code-switching for decades (see Deuchar, 2020; Gardner-Chloros, 2009, for overviews). This rich line of research has uncovered that code-switching is systematic (i.e., rule-governed) and is subject to the influence of various linguistic and extralinguistic factors such as grammatical category, congruency across the two languages, frequency of use, bilingual proficiency, and community use. Psycholinguists have more recently begun to explore the cognitive and neural mechanisms involved in mixed language use. This literature can be divided into two general approaches. Researchers primarily interested in bilingual language control and lexical access have resorted

https://doi.org/10.1075/bpa.12.07tom

to studying mixed language input in cued language switching paradigms (e.g., Meuter & Allport, 1999; Gollan & Ferreira, 2009) or by including artificial (i.e., unattested) switches in "connected speech" (e.g., Gollan & Goldrick, 2018; Schotter et al., 2019). Other researchers focus on successful production and comprehension of code-switching as reflective of bilingual language use (see Beatty-Martínez et al., 2018; Valdés Kroff et al., 2018; Van Hell et al., 2018, for overviews). Code-switching is speaker-generated in production, embedded within sentential contexts, and can occur at various syntactic junctures. In contrast, language switching is externally cued (via background color or auditory cue) and most typically studied with single lexical items, which come from a single grammatical category (i.e., nouns). In this chapter, we will focus on sentence processing in sentential code-switching as we are principally concerned with how bilingual comprehenders integrate code-switches in real-time processing. Our goal is to unravel how what on the surface appears "costly" may in fact, show processing benefits in certain contexts.

In broad terms, psycholinguistic approaches to code-switching demonstrate that bilinguals experience greater processing costs when encountering a code-switch as compared to non-switched utterances (e.g., Altarriba et al., 1996), a phenomenon referred to as "switch costs." However, a distinction should be made between switch costs in production and integration costs in comprehension. Within the cued language switching literature, switch costs typically refer to greater naming latencies when switching between languages at a trial-by-trial basis. Bilinguals experience switch costs in production (e.g., Meuter & Allport, 1999) likely due to the exogenously cued locus of the experimental paradigm but show reduced to no integration costs in comprehension (Declerck et al., 2019). In contrast, in studies that focus on sentential code-switching, bilinguals may benefit and show reduced switch costs in production (Beatty-Martínez et al., 2020), likely due to speakers exploiting speech planning mechanisms, but instead show costs to integration in certain contexts (Altarriba et al., 1996).

While sentential code-switch integration costs have been robustly documented in the prior literature, recent approaches have attempted to find contexts in which these can be reduced. These studies demonstrate that integration costs are attenuated by linguistic factors such as syntactic distribution or phonetic cues (Beatty-Martínez & Dussias, 2017; Fricke et al., 2016; Guzzardo Tamargo et al., 2016), phonotactic constraints (Li, 1996), cognitive mechanisms such as syntactic priming (e.g., Kootstra et al., 2010, 2012) or lexical triggering (Broersma & de Bot, 2006), individual-level and task factors such as proficiency, switch direction, proportion of code-switches (Johns et al., 2019; Litcofsky & Van Hell, 2017), social cues (Kaan et al., 2020; Valdés Kroff et al., 2018), or the frequency of switching within a community (Adamou & Shen, 2019). These approaches have in some cases documented diminished costs to varying degrees, but do not find evidence for eliminated integration costs. On the surface, these robust processing costs (even if diminished) present an unusual paradox for bilingual language use and prediction (Altmann & Kamide, 1999). For one, bilingual code-switching is ubiquitous within bilingual communities, yet these lab-based results would suggest that bilinguals are engaging in effortful processing. This paradox may be the result of the artificial setting and decontextualized stimuli in which lab-based studies on code-switching are carried out, essentially turning psycholinguistic studies on code-switching into investigations on how bilinguals process unexpected input (Gullberg et al., 2009; Moreno et al., 2002; Valdés Kroff et al., 2018). Here, we will propose an alternative approach as it relates to predictive processing by shifting our focus towards potential benefits the code-switch provides to downstream processing.

Prediction, in our view, consists of the probabilistic activation of upcoming lexical, semantic, and grammatical information (Kuperberg & Jaeger, 2016). In unilingual processing, incremental, predictive processing leads listeners and readers to probabilistically anticipate upcoming information by continuously updating their expectations for upcoming information based on prior context (e.g., Altmann & Kamide, 1999). Thus, individuals are anticipating grammatical structure, semantic concepts, and potentially lexical forms. In appropriate pragmatic contexts (i.e., in the presence of other known bilinguals), the bilingual is additionally attempting to anticipate possible code-switches, as shown by studies demonstrating bilinguals' sensitivity to cues preceding code-switches (Beatty-Martinez & Dussias, 2017; Shen et al., 2020). So why do bilinguals code-switch if it is costly or difficult to predict? This paradox leads us to argue for a different approach to investigate prediction in code-switching. Our proposal is to shift focus from the integration costs at the site of the code-switch itself and ask how code-switches affect processing and prediction downstream. In other words, even if bilinguals experience measurable integration costs when they first encounter a code-switch, these costs may later turn into processing benefits because the code-switch serves as a contextual signal to the bilingual comprehender. This proposal is based on socio-pragmatic and information-distribution functions that have been linked to the production of code-switching. For example, bilinguals produce code-switches before more difficult or less expected content (Myslín & Levy, 2015; Example (1)) and before the introduction of socially negative, taboo topics (Bentahila, 1983; Tomić, 2015; Example (2)).

(1) a. Tady vidiš že ona je *in need*.
'Here you see that she is *in need*.'
b. A potřebuje *entertainment*.
'And she needs *entertainment*.' (Myslín & Levy, 2015, p. 872)

In two consecutive utterances in Example (1), adapted from Myslín & Levy (2015), the authors argue that newly introduced concepts, such as the concept of NEED, italicized in a., are expressed in English. When the concepts become discourse-old, they are expressed in Czech, such as *potřebuje* ('needs') in b.

(2) J: neću da mi to radiš
'I don't want you to do that to me'
L: dobro, za početak ću da podrigujem
'ok, for starters I'll burp'
J: you know I'm gonna start *farting* as a {trails off}
L: exactly, that was my goal all along (Tomić, 2015, p. 353)

In Example (2), adapted from Tomić (2015), the author observes a speaker of Serbian switching to their second language (L2) English, the language of power and global majority, to discuss taboo concepts (italicized).

These socio-pragmatic choices may signal a trade-off between immediate integration costs and downstream facilitated processing. Our experimental framework taps into this facilitative function for the bilingual comprehender by explicitly designing psycholinguistic experiments that manipulate the language that bilingual comprehenders encounter and examining whether processing a code-switch leads to better prediction or facilitated processing of upcoming linguistic input. In this chapter, we illustrate this approach with two studies where we summarize the basic concept, methods and design, and the primary results of interest. These illustrations are necessarily brief, and we include references to more detailed reports. We conclude our chapter by discussing new insights on prediction in bilingual sentence processing with the goal of shifting focus from integration costs into processing benefits. This shift starts from the premise that bilingual code-switching is a common bilingual language practice that underscores the highly adaptive bilingual comprehension system.

Illustrative Study 1: Can code-switching signal less expected upcoming lexical information?

Our first illustrative study tests the effects of processing code-switches in the auditory modality on the prediction of upcoming words based on their lexical frequency (Tomić & Valdés Kroff, 2021a). This study builds on Myslín and Levy (2015), who found that US Czech-English bilinguals switch to English on less predictable words, presumably to improve listeners' comprehension. This preferred switch direction to the language of power is found commonly in contact settings with a power asymmetry between languages (Blokzijl et al., 2017). The authors created a corpus of Czech-English bilingual discourse, containing Czech-only utterances and Czech to English code-switched sentences, where the code-switch occurred on the final word. They calculated a range of lexical accessibility and syntactic dependency factors for code-switched vs. unilingual final words in the utterances to ascertain the factors which most affect code-switching behavior. The main factor of interest was predictability of meaning in context, which was calculated using a Shannon guessing game, in which a separate pool of participants guessed the meaning of the utterance-final word. The predictability of meaning was defined as the percentage of correct guesses. Statistical modeling of how CS behavior was affected by these factors showed that less predictable words were frequent code-switch sites from the minority language, Czech, to the majority, more salient or marked language, English (Myslín & Levy, 2015, p. 878). The authors reason that this practice is due to audience design, i.e., speakers are aware of which language is considered more salient by their listeners and use it to encode more difficult meaning to ensure it is comprehended. Simply put, in the context of this study, the older Czech-English bilinguals are likelier to stay in Czech when talking to other bilinguals, unless signaling harder to process content. Czech is thus the less marked or salient language and English the more marked or salient language. If this information-theoretic function is indeed a general function in the pragmatics of code-switching, then switch direction and the salience of the languages involved in switching are likely agreed upon as a community practice and not based on the individual proficiency and language dominance of bilingual speakers (e.g., Bhatt & Bolonyai, 2011).

Our study uses the visual world paradigm, a common eye-tracking paradigm in which participants view a visual scene while listening to audio instructions (Tanenhaus et al., 1995), to test whether bilingual listeners exploit this code-switch distribution pattern during online processing. If so, then bilinguals should be able to predict less predictable words after a code-switch. We operationalized predictability through the correlated measure of lexical frequency (Calvillo et al., 2020). The study design took inspiration from visual world studies testing the effect of disfluent speech on prediction. Disfluent fillers, such as 'uhm', 'uh', are also found to precede unexpected, discourse-new words (Arnold et al., 2000; Barr, 2001). Consequently, listeners use this distributional information to adjust their expectations for upcoming information. In the case of disfluent vs. fluent speech, listeners expected less frequent or discourse-new targets (Arnold et al., 2003, 2004, 2007).

For the experimental design, we extracted a set of images illustrating more and less frequent words from the International Picture Naming Project (IPNP, Bates et al., 2003; Szekely et al., 2003, 2004, 2005). Each experimental panel contained two images, corresponding to a lower and a higher frequency noun respectively (Figure 1A), which could vary in their location on a computer screen (Figure 1B). We controlled for the frequency of the object label equivalents in Spanish and

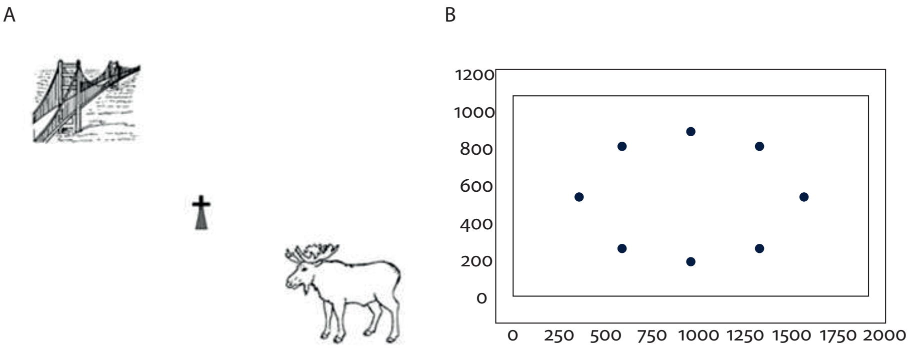

Figure 1. A. The illustration of an experimental picture panel for the "bridge" and "moose" high- and low-frequency label image pair. B. Possible positions for images

English and for the gender of the low- and high-frequency counterparts in Spanish. At the same time, we made sure to create significant frequency differences for each experimental pair as confirmed by a one-tailed paired *t*-test on experimental lists (see Tomić & Valdés Kroff, 2021a).

A highly proficient Puerto Rican Spanish-English bilingual recorded Spanish unilingual (Sp) and Spanish-English code-switched (CS) instructions to click on an image:

(3) Sp: *Encuentra el dibujo de un/una/Ø* ___________
CS: *Encuentra el* drawing of a/an/Ø ___________
'Find the drawing of a/an/Ø ___________'
Sp: *Elige el dibujo de un/una/Ø* ___________
CS: *Elige el* drawing of a/an/Ø ___________
'Select the drawing of a/an/Ø ___________' (Tomić & Valdés Kroff, 2021a)

We offset the code-switch away from the target by several words to avoid immediate effects of integration costs on the target word and test whether code-switching serves as a predictive cue for upcoming linguistic content. The code-switched instructions were naturally pronounced with a slight prolongation before the CS onset compared to the Spanish version (mean difference = 22 ms). We left this delay unchanged, as slight delays have been found to precede code-switches in bilingual discourse and aid the processing of a CS (Fricke et al., 2016). Taking these cues out can make CS processing more difficult (Shen et al., 2020). We chose only the Spanish-English code-switching direction as it is the most frequent code-switching direction among Spanish-English bilinguals in the US (Blokzijl et al., 2017; Moreno et al., 2002; Valdés Kroff et al., 2018). This preferred switch direction tracks with Myslín and Levy's (2015) attested switches from a minority language (in this case

Spanish) into the majority language (i.e., English). Using the switch direction which is not attested to co-occur with less expected items would have likely introduced confounds. Similarly, to create the ecological context of bilingual code-switching, English unilingual sentences were not included, which also follows the experimental design of Myslín and Levy (2015).

Thirty Spanish-English bilinguals who regularly code-switch listened to auditory instructions and clicked on correct images via computer mouse. They were exposed to 8 pairs of images in each condition (frequency [low, high] × language of instructions [Spanish, code-switches]) in 32 experimental visual scenes, as well as 64 filler visual scenes (accessible through Open Science Framework [OSF] repository https://osf.io/azcn4). Filler visual scenes had the exact same structure as experimental trials, yet they included mostly mid-range frequency items. Also, the filler pairs of images did not significantly differ in frequency and the instructions accompanying them were always in Spanish. The task was preceded by experiment procedure instructions containing code-switches to promote a bilingual language mode and the global expectation that code-switches may occur during the experimental session. Language proficiency tests, questionnaires, and self-reported proficiency measures showed that these bilinguals overall acquired Spanish first but were currently dominant in English (full details in Tomić & Valdés Kroff, 2021a). We conducted a Growth Curve Analysis (Mirman, 2014) on the proportion of looks to images in the time period from 200 ms before to 200 ms after the target onset. Eye movements in this time span reflect predictive processing, before the participants' eye movements are affected by the target word processing, as planned eye movements generally take 150–200 ms to launch (Allopenna et al., 1998; Travis, 1936). Thus, we are focusing on the impact that listening to a code-switch has on predictive processing of upcoming lexical information. In the remainder of this section, we summarize the main findings from this study. For full model results, see Tomić & Valdés Kroff (2021a) and the associated OSF repository <https://osf.io/azcn4>.

The statistical model reveals two important interactions: language of instructions × frequency and language of instructions × frequency × language dominance. The critical, first interaction indicates that bilinguals looked at low frequency images significantly more in the code-switched condition than in the non-switched Spanish condition (Figure 2), suggesting that bilinguals interpreted the code-switch as a contextual signal for upcoming, less expected lexical content.

The second interaction further reveals that eye-movement patterns were modulated by language dominance, such that participants more dominant in Spanish exhibited the typical frequency bias towards high frequency nouns in the non-switched Spanish condition, i.e. they looked more at objects representing higher frequency words (Dahan et al., 2001). In contrast, the less Spanish-dominant participants did not show a clear frequency bias in Spanish-only conditions. Nevertheless, Spanish

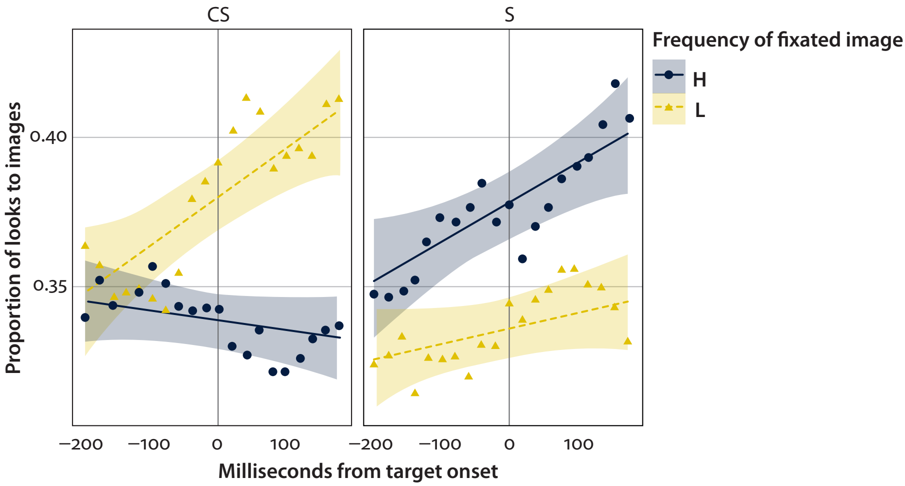

Figure 2. Proportion of looks to images with Low (L) and High (H) frequency labels 200 ms before and after the target word onset (vertical line), split by language condition: Code-switched (CS) and Spanish (S); adapted from Tomić & Valdés Kroff (2021a)

dominance did not affect the increased looks to low-frequency items in the CS condition, indicating that participants interpreted code-switches as contextual and predictive cues in sentence processing. These results suggest that Spanish-English bilingual listeners are exposed and sensitive to the distribution pattern of code-switches to the language of power as a pragmatic function of encoding less expected upcoming information, regardless of their own individual language dominance (Tomić & Valdés Kroff, 2021a). The results thus corroborate the audience-design interpretation that Myslín and Levy (2015) propose for code-switching.

This study is the first to experimentally test and confirm the influence of code-switches on the predictive processing of upcoming linguistic content. Specifically, processing a code-switch helps bilinguals anticipate upcoming less expected information, operationalized using lower frequency items, due to bilinguals being exposed to the pattern of code-switches preceding or occurring on less expected items in production (Myslín & Levy, 2015). Unlike foreign accent (Romero-Rivas et al., 2016), or foreign accent combined with filled pauses (Bosker et al., 2014), which cause prediction to halt likely due to the lack of exposure to such linguistic contexts, we show that code-switches are interpretable as beneficial cues to lower-frequency information. Processing a code-switch steers bilingual prediction patterns away from a general, high-frequency heuristic to preactivate lower-frequency information (Dahan et al., 2001) and helps bilingual comprehenders predictively attend to lower-frequency words. We discuss possible mechanisms and the scope of this effect in the Discussion section.

Illustrative Study 2: Can code-switching ease the processing of taboo or negative information?

Having initial evidence of bilingual listeners interpreting code-switching as a facilitatory cue for upcoming speech, we turn to other co-occurrences of code-switches with different linguistic content in production to investigate whether code-switches lead to reduced integration costs for socially negative, taboo content (Tomić & Valdés Kroff, 2021b). Sociolinguists have observed that code-switches often precede embarrassing, taboo, and negative content (Bentahila, 1983; Bond & Lai, 1986; Tomić, 2015; Example (4)):

(4) waħed lli ʕandu *la diarrhée* tajSwb šwija
'someone who has diarrhea can take a bit of it'
(Arabic-French, Bentahila, 1983, p. 236)

Following the logic of Study 1, this socio-pragmatic function of code-switching in production led us to hypothesize that code-switches can serve as a facilitatory cue of upcoming emotional taboo words in comprehension.

Emotional reactivity or emotionality to words and other stimuli has been captured by two main dimensions, *arousal*, how exciting the stimulus is, and *valence*, how positive or negative it is perceived to be (Bradley et al., 2001; Osgood et al., 1957). In monolingual studies, positive words are processed faster and/or responded to more accurately than neutral words and often negative words (Kissler & Koessler, 2011; Schacht & Sommer, 2009). The positive word behavioral advantage has been termed "positivity bias" (Herbert et al., 2009). Negative words sometimes also show facilitation (Knickerbocker et al., 2015), but often only in early processing or modulated by other factors, such as arousal (Hofmann et al., 2009) and lexical frequency (Scott et al., 2012; Scott et al., 2014). Consequently, negative and taboo words have been hypothesized to grab attention more intensely, due to negative stimuli being more pertinent to survival than positive ones (Baumeister et al., 2001; Mackay et al., 2004; Pratto & John, 1991), even when the task does not require attending to valence or tabooness (Eilola & Havelka, 2011; Pratto & John, 1991). This increased attentional demand has been interpreted as the locus for slowdowns in behavioral measures (negative: Eilola & Havelka, 2011; Pratto & John, 1991; Schacht & Sommer, 2009; taboo: Eilola & Havelka, 2011; Raizen et al., 2015), often in later processing measures reflective of integration (Lüdtke & Jacobs, 2015).

As for bilinguals, initial evidence indicated that they do not develop emotional reactivity in their L2 to the extent of their first language (L1; e.g. Anooshian & Hertel, 1994). Nevertheless, recent evidence using different experimental techniques, including eye-tracking (Sheikh & Titone, 2016) and event-related potentials (ERPs, Opitz & Degner, 2012), confirm that bilinguals do develop emotional reactivity in their L2 at a similar level to their L1, modulated by individual-level factors such as increased proficiency and naturalistic exposure to and use of the language (Altarriba & Basnight-Brown, 2011; Ayçiçeği & Harris, 2004; Conrad et al., 2011; Opitz & Degner, 2012; Ponari et al., 2015). However, none of these prior bilingual emotionality studies directly examines the emotional reactivity of bilinguals in code-switched discourse, despite the relative frequency of CS in bilingual discourse (~20%, Beatty-Martinez & Dussias, 2017) and the aforementioned socio-pragmatic function of CS to encode upcoming negative or taboo topics.

Oganian et al. (2016) tested the extent of decision bias, a supposed consequence of emotional reactivity, when a problem was presented in L1, L2, and in code-switched passages. One type of decision bias is the framing effect, referring to the tendency to choose a particular option based on the positive (gains) vs. negative (loss) framing of the problem (Oganian et al., 2016). The authors found that the framing effect was significantly reduced in the code-switched condition but not in the L2, in line with prior sociolinguistic work suggesting that CS could modulate emotional reactivity. Our study tests this hypothesis using eye-tracking-while-reading measures as direct correlates of emotional reactivity. Emotionality has been shown to affect both early and late eye-tracking measures, such that more emotional words, especially positive words, are processed faster (Knickerbocker et al., 2015; Scott et al., 2012; Sheikh & Titone, 2013). The emotionality effects for words in general, including in eye-tracking studies, have been explained by the motivational salience of emotional words, which promotes their accessibility and/or recruits additional resources to enhance their processing (Mackay et al., 2004). As a signature of emotional reactivity, we can expect that positive words are read faster than neutral words. Negative and taboo words could be initially read faster, but cause delays in processing in later eye-tracking measures, due to additional resources deployed to attend to their processing (Eilola & Havelka, 2011; Pratto & John, 1991; Schacht & Sommer, 2009; Raizen et al., 2015).

We designed an eye-tracking-while-reading experiment to examine the effects of code-switches on the processing of taboo words in an initial 2 × 3 experiment design: tabooness (taboo, neutral) × language (English, Spanish, code-switched). We embedded taboo and neutral words in unilingual Spanish and English sentences and Spanish-English code-switched sentences (Table 1; materials available through the OSF repository: https://osf.io/du7ay/). As in the case of Study 1, we chose the Spanish to English code-switch direction since switches to the language of power, usually the local and/or global majority language, are more frequently attested in bilingual communities (e.g., Nicaraguan English Creole to Spanish in Nicaragua: Blokzijl et al., 2017; Bhatt on languages in India, 2013, as cited in Blokzijl et al., 2017; Spanish to English in the US: Blokzijl et al., 2017; Poplack, 2000; Zentella, 1997). Also, sociolinguistic studies indicate that switches prior to or on taboo words are from the L1 to the L2 or from the less to the more situationally marked language (Bentahila, 1983; Bond & Lai, 1986; Tomić, 2015). English, in US Spanish-English bilingual communities, is both the local and the global majority language, as well as the language of power, thus making it the situationally marked language during minority language use.

Table 1. Experimental design for Illustrative Study 2

Lang	Status	Item
CS	Taboo	A principios de este verano, they found a *turd* in the showers at the waterpark.
CS	Neut	A principios de este verano, they found a *tooth* in the showers at the waterpark.
Eng	Taboo	Earlier this summer, they found a *turd* in the showers at the waterpark.
Eng	Neut	Earlier this summer, they found a *tooth* in the showers at the waterpark.
Sp	Taboo	A principios de este verano, encontraron una *cagada* en las duchas del parque acuático.
Sp	Neut	A principios de este verano, encontraron un *diente* en las duchas del parque acuático.

Target words italicized. Sp = Spanish, Eng = English, CS = code-switched

We devised 48 target neutral-taboo word pairs (46 noun pairs and 2 adjective pairs), in English and Spanish, e.g. 'cocksucker' – *chupapollas*, 'airplane' – *avión*. The taboo words denoted body parts used for copulation and excretion, excrement matter, and sexual acts. Neutral words represented non-taboo words which could semantically and syntactically fit the same sentence frame. We pre-normed the target words on the dimensions of use and exposure to words, arousal and valence, as defined above, offensiveness (how offensive the word is to the person), tabooness (how taboo the word is in society), and imagery, following Janschewitz (2008). Taboo words were rated significantly lower on exposure, use, valence, and higher on offensiveness, tabooness, and arousal (full report in Tomić & Valdés Kroff, 2021b).

To maximize ecological validity, the sentences were conversation-like and featured code-switches deemed natural by two proficient Spanish-English bilinguals. Two to three words intervened between the CS and target word, as in Study 1, to avoid immediate integration costs and to investigate whether the CS leads to subsequent predictive processing. We additionally conducted cloze probability and plausibility pre-norming studies on the sentence frames and combinations of the sentence frames and target words (Sheikh & Titone, 2013). With the cloze probability pre-norming task, we confirmed the unpredictability of target words based on pre-target sentential context. The Plausibility pre-norming task included participants judging how likely it would be for the taboo and neutral versions of Spanish and English sentences to be uttered in daily life. The linear mixed effects models for Spanish and English separately showed that taboo versions of sentences were less likely in daily conversations. The low plausibility of taboo words is an intrinsic property of taboo words. Therefore, we did not control for this variable.

We created 6 experimental lists with 48 experimental sentences and 8 items per condition. The study was split into a Spanish and CS block (32 experimental sentences) and an English monolingual block (16 experimental sentences) to mimic the language contexts in which bilinguals find themselves during their daily lives in the US (Zentella, 1990). We included additional 96 neutral filler sentences. The filler sentences for the Spanish and CS block included 56 Spanish and 8 code-switched sentences, to keep the ratio of CS utterances close to their attested distribution in production (Beatty-Martinez & Dussias, 2017).

Thirty Spanish-English bilinguals, who acquired both languages before the age of 12 and reported code-switching regularly, participated in the experiment. Participants completed several measures to establish a more encompassing language and emotionality profile: adapted standardized tests for Spanish (Diplomas of Spanish as a Foreign Language, DELE, Ministry of Education, Culture, and Sport of Spain, 2006) and English (Michigan English Language Institute College English Test, MELICET, University of Michigan English Language Institute, 2006), Language History Questionnaire (LHQ, Guzzardo Tamargo et al., 2016), the Autism Quotient (Baron-Cohen et al., 2001), and the Emotionality Questionnaire (EQ, Janschewitz, 2008) at the end of the experiment. The EQ required participants to rate the experimental words they read as in the pre-norming study, on exposure, use, valence, arousal, tabooness, offensiveness, and imagery. These ratings were used as measures of emotionality in the final statistical analysis, to control for the language variant and individual rating variability.

The participants overall acquired Spanish first. Nevertheless, as in Study 1, the bilinguals were more dominant in English, as demonstrated by their proficiency scores and self-reports (full report in Tomić & Valdés Kroff, 2021b). Additionally, the participants were overall more exposed to English in everyday life, and they reported that they would address a fellow bilingual in English most of the time. Order and age of acquisition (Pavlenko, 2012), proficiency, language dominance, and naturalistic exposure (Altarriba & Basnight-Brown, 2011; Degner et al., 2012) are important factors for the development of emotionality. In our study sample, Spanish had a slight advantage in terms of order and age of acquisition, yet English had the advantage due to increased exposure and dominance. Importantly, 28 participants out of 30 confirmed they code-switch in the LHQ questionnaire. No participants scored higher than 32 points on the AQ questionnaire, the cut-off for the clinical diagnosis of autistic traits (Baron-Cohen et al., 2001). The models with the AQ scores as predictors did not provide a significant improvement, so we do not report them here.

Participants completed the experiment in two blocks on separate days. The instructions were presented in code-switched speech for the Spanish and Code-switched block, and only in English for the English block, to activate the appropriate language schemas. Participants read the sentences and responded to comprehension questions.

We fit linear mixed-effects models using the lme4 package (Bates et al., 2015) in R (R Core Team, 2017) to Gaze Duration (GD) and Total Duration (TD) eye-tracking measures for the target and post-target regions. GD is the sum of all fixations within a region before eyes move out of the region. TD is the sum of all fixations on the region, including regressions. The post-target region consisted of two words following the target word in order to capture any spillover effects.

The study was designed to investigate the interaction of tabooness and language. Nevertheless, initial models with tabooness as the only emotionality variable did not show any significant effects. We suspected that this was due to the target words not being well controlled for valence, i.e., experimental word pairs were not matched for valence and both neutral and taboo target words could be positive or negative. Valence and arousal primarily and separately contribute to emotional reactivity (see for a review Citron, 2012), and there is evidence that positive and negative words exhibit different behavioral effects (Kissler & Koessler, 2011). Since the experimental words, including neutral and taboo words, included an approximately equal number of positive and negative words as judged by participants and since arousal and valence were not correlated for the target region, we added valence and arousal ratings provided by the participants in the final statistical models. Tabooness was not included in the final models, as it was heavily correlated with both arousal and valence.

As the measure of significance, we compared the base linear mixed effects models for each measure and region to models including incrementally added emotionality variables of valence and arousal. The base linear mixed effects models included language (English, Spanish, CS [reference level]), Spanish dominance (DELE divided by MELICET scores), Word Length (character length of the target or the post-target region), and their interactions as fixed effects, and minimally Participant and Item random intercepts. To this base model, we added emotionality factors as fixed effects, as well as interactions with other variables, to determine the scope of emotionality effects. All continuous variables were normalized and the dependent variables were trimmed and transformed following adapted standard procedures (e.g. Sheikh & Titone, 2013). For GD, the addition of emotionality variables either did not improve the model fits, or it did not produce any significant language × emotionality interactions. In the remainder of this section, we discuss the relevant language × emotionality (arousal and valence) interactions from the TD models for the target and spillover region. Full model reports, including additional analyses, can be found in the associated OSF repository <https://osf.io/du7ay>.

In the summarized models, there was a main effect of language (Spanish), such that target words in Spanish were read slower than the words in the CS condition. This is likely due to the fact that the CS condition contained words in English, most participants' dominant language. Heritage bilinguals also might not have developed reading skills in Spanish to the same degree as in English. The language (English) × valence interaction was trending in the TD model for the target region, such that more positive words were read faster in English than in the CS condition. This trending interaction suggests facilitation in English for positive words due to emotional reactivity, as well as emotionality reduction in the CS condition. Increasing Spanish dominance led to a smaller TD difference between negative and positive words for unilingual conditions (flattened slopes for English and Spanish in Figure 3, right panel, compared to left panel). For these more Spanish-dominant participants, more negative words were read slower and more positive words faster in the CS condition (solid line, Figure 3, right), resembling the slopes of the emotionality effect in unilingual conditions for the less Spanish dominant speakers (dashed and dotted lines, Figure 3, left). This was corroborated by the negative-coefficient two-way interaction of valence and Spanish dominance, with CS as the language

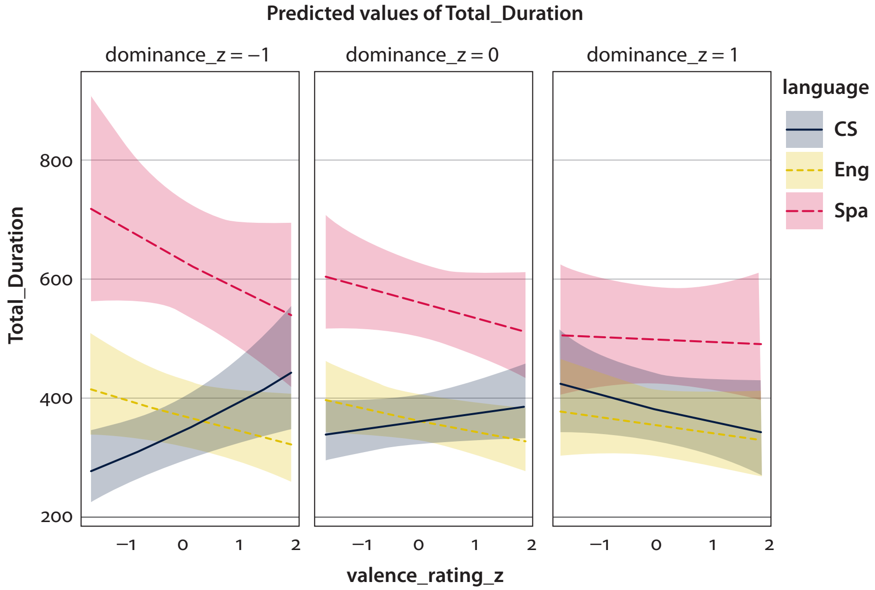

Figure 3. Predicted marginal effects plot for the language × valence x dominance interaction for the reported TD model for the target region. Dominance rating of $z = -1$ corresponds to more English-dominant speakers. Valence rating of $z = -1$ corresponds to negative valence words

baseline, indicating that positive words were read faster in the CS condition for more Spanish dominant participants. These two interactions suggest that more Spanish-dominant participants do not experience the same emotionality reduction in the CS condition compared to the less Spanish-dominant speakers (compare solid lines across Figure 3). This and subsequent predicted marginal effects plots were constructed using the ggeffects R package (Lüdecke, 2018).

With an increase in word length, more Spanish dominant participants read more arousing words slower in the English vs. CS condition. Predicted marginal effects for this interaction while keeping Spanish dominance at maximum suggests that Spanish dominant participants do not experience facilitation for arousing words in English, particularly when the words are long.

The most parsimonious model for the TD measure in the spillover region included arousal and valence, and their interactions with other variables. For the relevant language and emotionality interactions, spillover words were read faster in English the more positive the target words were, whereas this reading pattern was reversed in the CS condition (Figure 4), suggesting that code-switches reversed the emotionality signature from the English unilingual condition. This interaction mirrors the trending interaction in the TD model for the target region. With increasing

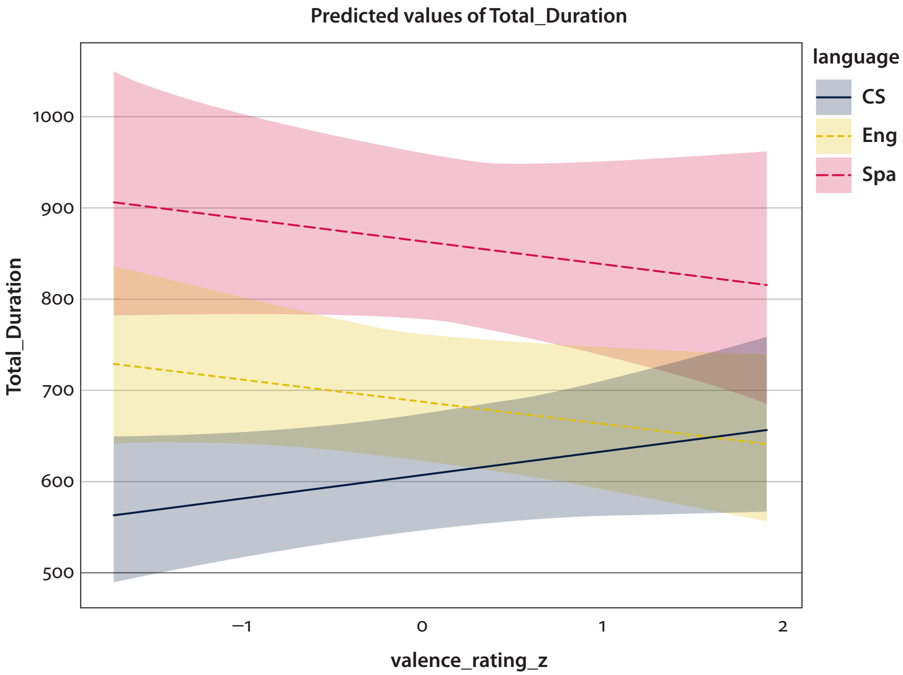

Figure 4. Predicted marginal effects plot for the language × valence interaction for the reported TD model for the post-target region. Valence rating of $z = -1$ corresponds to negative valence words

Spanish dominance, the spillover words after more positive and more arousing target words were read slower in English as opposed to CS sentences.

Research on emotionality shows that emotional words exhibit facilitated processing, at least definitively so for positive words (Kissler & Koessler, 2011; Schacht & Sommer, 2009). Negative words provide mixed results, behaving similarly as neutral words, or causing processing slowdowns (Kissler & Koessler, 2011; Pratto & John, 1991; Schacht & Sommer, 2009). Taboo words consistently cause slowdowns in processing in behavioral studies (e.g., Raizen et al., 2015). Therefore, the facilitation we see for positive words in the English condition compared to the CS condition is likely the manifestation of strong emotional reactivity in English and its significant reduction after a code-switch, despite the fact that the CS and English language conditions feature exactly the same English words. We did not, however, find a strong emotionality effect in Spanish. This asymmetry is likely due to more variable Spanish proficiency and emotionality, caused by the groups' status as heritage speakers of Spanish. In the case of code-switched trials, the negative coefficient of the language (English) × valence interaction, trending in the target region and significant in the post-target region, suggests that preceding CS facilitates the processing of linguistic input following negative words, unlike in English unilingual processing. This is likely caused by the pre-activation or facilitated integration of negative content, a specific valence value, after a code-switch, due to the attested pattern of code-switches preceding negative words. When these expectations are not met, the participant likely experiences prediction error and adjusts their expectations on the type of content likely to follow a CS. More generally, the results show that bilinguals build expectations for upcoming linguistic input based on encountering a code-switch.

The type of information which is pre-activated in Study 2 might be related to Study 1. Negative emotional words are generally more informative and less frequent than positive words in discourse (e.g., Boucher & Osgood, 1969). Therefore, participants might expect generally more informative, less frequent content after a CS, including low frequency words *and* negative words as a subset of low frequency words. Interestingly, Example (1), which Myslín & Levy (2015) give as an illustration of less predictable words being switched to English, could also be analyzed in terms of the CS pattern of preceding or occuring on negative social concepts. From the authors' explanation, both code-switched expressions, *in need* and *entertainment,* are related to the behavior the speaker finds promiscuous and socially negative.

As in Study 1, (Spanish) dominance played a role in this overall emotionality effect. Nevertheless, in Study 2 increased Spanish dominance or more balanced proficiency attenuated the CS emotionality reduction effect. We discuss reasons for this modulation in the General Discussion section.

General discussion

With two recent studies, we have illustrated a novel approach in CS processing research, involving the investigation of the effects of bilingual CS on the (predictive) processing of upcoming linguistic content in terms of lexical frequency and emotional valence. These studies contribute to research on prediction, i.e., predictive pre-activation and/or facilitation (Kuperberg & Jaeger, 2016) of specific lexical input by adding another useful predictive cue, formerly characterized as a processing burden. We show that a code-switch could cause pre-activation of lower-level lexical features such as lexical frequency and emotional valence, which can be observed prior to target word onset (Study 1), or facilitate the processing of predicted information, even in the absence of an effect in early processing measures (GD; Study 2). Our studies do not speak directly to whether CS preactivates particular words, as the actual targets were unpredictable in both studies. Rather, the CS seems to be preactivating low-frequency information and negative valence, thus probabilistically constraining the pool of possible upcoming words to those that share these features and aiding their processing (DeLong et al., 2005).

The two purported types of preactivated information, lower lexical frequency and negative emotional valence, are related in several ways. There is a documented positivity bias in discourse, making positive words more frequent and negative valence or negative social concepts in general less frequent (Dodds et al., 2015). The negative valence could thus be a subcategory of low frequency information. The positivity bias in bilingual discourse in terms of frequency of positive concepts could be underlying the facilitative effects of positive words, also termed positivity bias, similar to the effects of the high-frequency bias (Dahan et al., 2001). Therefore, in both cases, the documented and potentially related biases of positivity and high frequency are overturned by encountering a code-switch.

It remains an open question as to which aspect of CS is responsible for adjusting prediction in such a way, that is, at which level of representation the information on the CS is encoded. The CS as a predictive cue could be richly represented as a language alternation, with potential information on the syntactic locus and/or specific direction of the language switch. If so, unattested or infrequent switches, e.g., from English to Spanish, could have caused the bilingual comprehender to stop predicting or adjust prediction in a different way. Nevertheless, the CS itself is a discrete salient point in the communicative stream not necessarily tied to linguistic information. Thus, the CS could also be represented as a manifestation of a sudden change in the intensity or type of linguistic or even paralinguistic information at some level of representation, acting like an attention signal to redirect the comprehender's expectations. Along similar lines, switches in register, genre, intonation, pace, brightness or boldness in writing could produce a similar effect on predictive pre-activation.

Additionally, we could expect CS to exert the same beneficial effect on the predictive processing of cognitively demanding linguistic information at other levels of representation. Some examples include helping bilinguals adjust their expectations to preactivate morphological complexity, length, discourse-newness, or unpredictability, in the service of facilitating the processing of morphologically complex vs. simple words, long vs. short words, discourse-new vs. -old concepts, or less predictable vs. more predictable words in certain contexts. These other means of organizing discourse content may have been at play in the studies we describe here. Further studies delineating or controlling for these variables, especially predictability in context, should test different levels of representation at which the linguistic material is pre-activated due to the presence of a CS, to further elucidate the scope and nature of the effect.

Devising experiments taking other CS functions and patterns into account would be especially useful for understanding the scope of this facilitatory effect, both in production and comprehension. Sociolinguists enumerate a slew of functions underlying CS, such as topic shift, quotation, addressee specification, reiteration, message qualification, clarification, emphasis, and interjections (Gumperz, 1982, pp. 75–84). Similar to the functions which inspired the illustrative studies, these socio-pragmatic functions suggest that CS precedes important junctions in speech and/or difficult or unexpected linguistic elements. Consequently, code-switches preceding addressee specification might help comprehenders anticipate and prepare for their turn, streamlining the conversation (De Ruiter et al., 2006). Code-switches used for topic shifts, reiteration, clarification, or emphasis might draw comprehender's attention and adjust predictive processing biases, thus potentially improving the processing and retention of new information or the information the speaker deems important.

Potential mechanisms underlying the CS effect

Several alternative or complementary mechanisms may underlie the CS effect that leads to facilitated processing of upcoming linguistic information. In terms of production, these CS co-occurrence patterns could be grounded in production difficulty: The "difficult" planned message might be momentarily more available in one language vs. the other. CS may be associated with production biases such as "easy first" in which harder elements of speech are produced later in utterances. Johns and Steucke (2020) recently presented evidence in support of this hypothesis by finding that code-switches are more likely to occur towards the end of intonational phrases and that normalized speech rate becomes faster after code-switches by analyzing a

New Mexican Spanish-English bilingual corpus of spontaneous speech. The faster speech rate after the code-switch likely signifies that bilingual speakers implement a code-switch when speech becomes difficult, thus proactively relieving production pressures, or that the code-switch occurs as a consequence of impending difficulty and serves to reactively alleviate speech planning difficulties. To some extent, this account parallels a similar underlying interpretation of disfluencies (Arnold et al., 2003, 2004, 2007). In other words, speakers are more likely to produce disfluencies before upcoming more "difficult" concepts. Comprehenders in turn might be sensitive to these production pressures.

The comprehender's sensitivity to the CS signal might be relatively automatic, through language processing mechanisms proposed in models such as the P-chain (Dell & Chang, 2014) or other integrated accounts of production and comprehension (Pickering & Garrod, 2013; see also Ito & Pickering, this volume). According to these models, comprehension entails constant prediction and modeling of the speaker's production either as a means to guide error-based learning or through simulation.

Alternatively, there may be no such automatic connection between comprehension and prediction. The bilingual comprehender could separately infer that a CS is related to the speaker's production difficulty, which is not necessarily something they personally experience. Here, the bilingual comprehender would use top-down information to make inferences on the speaker's perceived state of mind. Studies have shown that bilinguals can similarly pre-activate particular languages or bilingual language context on the basis of top-down information on the speaker's language habits or knowledge (Kaan et al., 2020; Molnar et al., 2015).

Another account is based on the use of the accumulated statistical distributions of how CS is produced. Whatever the cause of CS patterns in production may be, bilingual comprehenders may be tracking these statistical distribution patterns to guide comprehension (Production-Distribution-Comprehension model; MacDonald, 2013). These explanations have also been proposed for the effect of certain disfluencies on prediction (Arnold et al., 2003, 2004, 2007).

Ultimately, code-switching in production may serve several functions, including as a repair strategy for production difficulties or as a somewhat unlikely intentional speech event to direct the attention of the comprehender to upcoming salient points in speech. The contributions of these different CS functions in production could change with more frequent code-switching use or proficiency. Exposure to code-switching patterns and practices, proficiency, and other factors could also modulate bilingual comprehenders' use of code-switches for prediction. Future studies varying the social context (e.g. comprehender's knowledge on speaker's language use preferences or state of mind), production pressures, and methods (e.g., including ERPs) should be conducted to determine the potential role of these production and comprehension mechanisms in predictive CS effects.

Factors modulating how CS affects prediction

The degree to which a code-switch pre-activates upcoming lexical form likely depends on a number of factors tied to individual differences in language experience and cognition, the current socio-pragmatic context, and perceptual differences such as word length. Both studies demonstrate that language dominance plays a strong role, either relatively enhancing (Study 1) or diminishing (Study 2) the effects of code-switches on the prediction of low frequency and negative content, respectively. This effect could be directly related to proficiency and different pressures on processing in more and less dominant languages or how these individual-level factors interact with aspects of code-switching such as switch direction (Litcofsky & Van Hell, 2017). Predictive processing and emotional processing may be differentially affected by proficiency/dominance in each language, with emotionality potentially being more dependent on global emotionality levels in each language.

Finally, it is important to highlight that not all bilinguals code-switch or are exposed to code-switching in a uniform way. The apparent effect of dominance may simply be a manifestation of the linguistic knowledge of a bilingual due to use and/or exposure, where the more balanced or more Spanish-dominant participants in our study could also be less exposed to code-switching or this particular switching function. Such bilinguals might thus have difficulties integrating the switch itself and stop predicting, similar to the effects of foreign accent on prediction (Bosker et al., 2014). Due to the relatively uniform CS frequency of use and exposure measures in our participant samples, we could not investigate further how these factors affect the interpretation of CS as a predictive cue. Moreover, the ability to interpret CS as a predictive cue likely depends heavily on the comprehender's knowledge of the CS habits/patterns of a specific community or interlocutor (e.g., Adamou & Shen, 2019). Consequently, a bilingual listener may fail to interpret CS as a predictive cue in situations in which code-switches were not expected. Moreover, the typology of code-switching may affect bilingual comprehenders' ability to interpret a CS as predictive cue. For example, code-switches may be classified as insertional, alternational, or as congruent lexicalizations according to Muysken (2000). The dominant CS types that emerge in a community are subject to socio-historical factors, which in turn, likely affect bilingual comprehension. Such outcomes are captured by recent models of bilingualism that incorporate the variability of bilingual interactional contexts (e.g., Green & Abutalebi, 2013; Green & Wei, 2014). Our studies show that investigating CS in relatively meaningful sentences is sufficient to provide the appropriate context to affect prediction. Nevertheless, these studies only represent a beginning in understanding how to incorporate additional sociolinguistic variables to determine the scope of predictive processing in code-switching.

Conclusions

Despite the challenge of studying code-switch processing in more ecologically valid paradigms, incorporating sociolinguistically-informed observations into the experimental study of CS is paramount for constructing a sound psycholinguistic theory of bilingual sentence processing (Myers-Scotton, 2006). One means for accomplishing this goal is to shift focus from purely examining integration costs found at the moment of processing code-switches to investigating whether bilinguals infer the code-switch as a meaningful and beneficial signal. As the results from the illustrative studies described in this chapter suggest, bilinguals can do so, indicating that processing costs (i.e., integration costs) can turn into processing benefits (i.e., predictive processing). This approach represents a novel lens through which we can study the linguistic resources that bilinguals deploy in both production and comprehension and which cannot be studied in monolingual or unilingual contexts alone. Our approach here is promissory with room for additional discoveries that relate to prediction, emotionality, production planning, and comprehension.

Acknowledgments

We would like to thank Dr. Kathy Carbonell for generous assistance with stimuli recording for Illustrative Study 1. We would also like to thank Dhara Patel for her assistance with stimuli creation for Illustrative Study 2, as well as Atharva Chopde and Katelyn Porter for assisting with the data collection for Illustrative Study 1, and Wen Jiang, Jeffrey Kunath, and Alexis Berles, for assisting with the data collection for Illustrative Study 2.

Funding

Research was funded by the National Science Foundation Doctoral Dissertation Improvement Grant (BCS-1941514) to A.T and J.V.K and a Humanities Enhancement Scholarship to J.V.K.

References

Adamou, E., & Shen, X. R. (2019). There are no language switching costs when codeswitching is frequent. *International Journal of Bilingualism*, 23(1), 53–70. https://doi.org/10.1177/1367006917709094

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38(4), 419–439. https://doi.org/10.1006/jmla.1997.2558

Altarriba, J., & Basnight-Brown, D. M. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism*, 15(3), 310–328. https://doi.org/10.1177/1367006910379261

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, 24(4), 477–492. https://doi.org/10.3758/BF03200936

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition & Emotion*, 8(6), 503–514. https://doi.org/10.1080/02699939408408956

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*, 32(1), 25–36. https://doi.org/10.1023/A:1021980931292

Arnold, J. E., Kam, C. L. H., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 914–930. https://doi.org/10.1037/0278-7393.33.5.914

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science*, 15(9), 578–582. https://doi.org/10.1111/j.0956-7976.2004.00723.x

Arnold, J. E., Wasow, T., Ginstrom, R., & Losongco, T. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1), 28–55. https://doi.org/10.1353/lan.2000.0045

Ayçiçeği, A., & Harris, C. (2004). Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*, 18(7), 977–987. https://doi.org/10.1080/02699930341000301

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. https://doi.org/10.1023/A:1005653411471

Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In C. Cavé, I. Guaïtella, & S. Santi (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication* [Orality and gestuality: Multimodal interaction and behavior in communication] (pp. 597–600). L'Harmattan.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, 10(2), 344–380. https://doi.org/10.3758/BF03196494

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323. https://doi.org/10.1037/1089-2680.5.4.323

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: a bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology*, 11, 1699. https://doi.org/10.3389/fpsyg.2020.01699

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, 95, 173–189. https://doi.org/10.1016/j.jml.2017.04.002

Beatty-Martínez, A. L., Valdés Kroff, J. R., & Dussias, P. E. (2018). From the field to the lab: A converging methods approach to the study of codeswitching. *Languages*, 8, 19. https://doi.org/10.3390/languages3020019

Bentahila, A. (1983). Motivations for code-switching among Arabic-French bilinguals in Morocco. *Language & Communication*, 3(3), 233–243. https://doi.org/10.1016/0271-5309(83)90003-4

Bhatt, R. M. & Bolonyai, A. (2011). Code-switching and the optimal grammar of bilingual language use. *Bilingualism: Language and Cognition*, 14, 522–546. https://doi.org/10.1017/S1366728910000295

Blokzijl, J., Deuchar, M., & Parafita Couto, M. C. (2017). Determiner asymmetry in mixed nominal constructions: The role of grammatical factors in data from Miami and Nicaragua. *Languages*, 2, 20. https://doi.org/10.3390/languages2040020

Bond, M. H., & Lai, T. M. (1986). Embarrassment and code-switching into a second language. *Journal of Social Psychology*, 126(2), 179–186.

Bosker, H. R., Quené, H., Sanders, T., & De Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language*, 75, 104–116. https://doi.org/10.1016/j.jml.2014.05.004

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, 1(3), 276. https://doi.org/10.1037/1528-3542.1.3.276

Broersma, M., & de Bot, K. (2006). Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, 9, 1–13. https://doi.org/10.1017/S1366728905002348

Calvillo, J., Fang, L., Cole, J., & Reitter, D. (2020, November). Surprisal predicts code-switching in Chinese-English bilingual text. In B. Webber, T. Cohn, Y. He, & Y. Liu (Eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 4029–4039). Stroudsburg PA: ACL. https://doi.org/10.18653/v1/2020.emnlp-main.330

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, 122(3), 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Conrad, M., Recio, G., & Jacobs, A. M. (2011). The time course of emotion effects in first and second language processing: A cross cultural ERP study with German–Spanish bilinguals. *Frontiers in Psychology*, 2, 351. https://doi.org/10.3389/fpsyg.2011.00351

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, 42(4), 317–367. https://doi.org/10.1006/cogp.2001.0750

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, 82, 515–535. https://doi.org/10.1353/lan.2006.0130

Declerck, M., Koch, I., Duñabeitia, J. A., Grainger, J., & Stephan, D. N. (2019). What absent switch costs and mixing costs during bilingual language comprehension can tell us about language control. *Journal of Experimental Psychology: Human Perception and Performance*, 45(6), 771–789. https://doi.org/10.1037/xhp0000627

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism: Language and Cognition*, 15(1), 181–189. https://doi.org/10.1017/S1366728911000095

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120394. https://doi.org/10.1098/rstb.2012.0394

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121. https://doi.org/10.1038/nn1504

Deuchar, M. (2020). Code-switching in linguistics: A position paper. *Languages*, 5(2), 22. https://doi.org/10.3390/languages5020022

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., … & Megerdoomian, K. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, 112(8), 2389–2394. https://doi.org/10.1073/pnas.1411678112

Eilola, T. M., & Havelka, J. (2011). Behavioural and physiological responses to the emotional and taboo Stroop tasks in native and non-native speakers of English. *International Journal of Bilingualism*, 15(3), 353–369. https://doi.org/10.1177/1367006910379263

Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. *Journal of Memory and Language*, 89, 110–137. https://doi.org/10.1016/j.jml.2015.10.001

Gardner-Chloros, P. (2009). *Code-switching*. Cambridge University Press. https://doi.org/10.1017/CBO9780511609787

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 640–665. https://doi.org/10.1037/a0014981

Gollan, T. H., & Goldrick, M. (2018). A switch is not a switch: Syntactically-driven bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(1), 143–156. https://doi.org/10.1037/xlm0000462

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, 29(4), 499–511. https://doi.org/10.1080/23273798.2014.882515

Gullberg, M., Indefrey, P., & Muysken, P. (2009). Research techniques for the study of code-switching. In B. E. Bullock, & J. A. Toribio (Eds.), *The Cambridge handbook on linguistic code-switching* (pp. 21–39). Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.003

Gumperz, J. J. (1982). *Discourse strategies* (Vol. 1). Cambridge University Press. https://doi.org/10.1017/CBO9780511611834

Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, 89, 138–161. https://doi.org/10.1016/j.jml.2015.12.002

Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., & Kissler, J. (2009). Amygdala activation during reading of emotional adjectives – an advantage for pleasant content. *Social Cognitive and Affective Neuroscience*, 4(1), 35–49. https://doi.org/10.1093/scan/nsn027

Hofmann, M. J., Kuchinke, L., Tamm, S., Võ, M. L., & Jacobs, A. M. (2009). Affective processing within 1/10th of a second: High arousal is necessary for early facilitative processing of negative but not positive words. *Cognitive, Affective, & Behavioral Neuroscience*, 9(4), 389–397. https://doi.org/10.3758/9.4.389

Janschewitz, K. (2008). Taboo, emotionally valenced, and emotionally neutral word norms. *Behavior Research Methods*, 40(4), 1065–1074. https://doi.org/10.3758/BRM.40.4.1065

Johns, M., & Steucke, J. (2020). *Is codeswitching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech* [Manuscript submitted for publication]. Department of Spanish, Italian and Portuguese, Pennsylvania State University.

Johns, M. A., Valdés Kroff, J. R., & Dussias, P. E. (2019). Mixing things up: How blocking and mixing affect the processing of codemixed sentences. *International Journal of Bilingualism*, 23(2), 584–611. https://doi.org/10.1177/1367006917752570

Kaan, E., Kheder, S., Kreidler, A., Tomić, A., & Valdés Kroff, J. R. (2020). Processing code-switches in the presence of others: An ERP study. *Frontiers in Psychology*, 11, 1288. https://doi.org/10.3389/fpsyg.2020.01288

Kissler, J., & Koessler, S. (2011). Emotionally positive stimuli facilitate lexical decisions – An ERP study. *Biological Psychology*, 86(3), 254–264. https://doi.org/10.1016/j.biopsycho.2010.12.006

Knickerbocker, H., Johnson, R. L., & Altarriba, J. (2015). Emotion effects during reading: Influence of an emotion target word on eye movements and processing. *Cognition and Emotion*, 29(5), 784–806. https://doi.org/10.1080/02699931.2014.938023

Kootstra, G. J., Van Hell, J. G., & Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language*, 63, 210–236. https://doi.org/10.1016/j.jml.2010.03.006

Kootstra, G. J., Van Hell, J. G., & Dijkstra, T. (2012). Priming of code-switches in sentences: The role of lexical repetition, cognates, and language proficiency. *Bilingualism: Language and Cognition*, 15(4), 797–819. https://doi.org/10.1017/S136672891100068X

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Li, P. (1996). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language*, 35, 757–774. https://doi.org/10.1006/jmla.1996.0039

Litcofsky, K. A., & Van Hell, J. G. (2017). Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia*, 97, 112–139. https://doi.org/10.1016/j.neuropsychologia.2017.02.002

Lüdecke, D. (2018). ggeffects: Tidy data frames of predicted marginal effects from regression models. *Journal of Open Source Software*, 3(26), 772. https://doi.org/10.21105/joss.00772

Lüdtke, J., & Jacobs, A. M. (2015). The emotion potential of simple sentences: additive or interactive effects of nouns and adjectives?. *Frontiers in Psychology*, 6, 1137. https://doi.org/10.3389/fpsyg.2015.01137

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, 226. https://doi.org/10.3389/fpsyg.2013.00226

Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, 32(3), 474–488. https://doi.org/10.3758/BF03195840

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Ministry of Education, Culture, and Sport of Spain. (2006). *Diploma de español como lengua extranjera* (Diploma of Spanish as a Second Language) Examination. Retrieved from http://www.dele.org/

Mirman, D. (2014). Growth curve analysis: A hands-on tutorial on using multilevel regression to analyze time course data. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36), 51–52.

Molnar, M., Ibáñez-Molina, A., & Carreiras, M. (2015). Interlocutor identity affects language activation in bilinguals. *Journal of Memory and Language*, 81, 91–104. https://doi.org/10.1016/j.jml.2015.01.002

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language*, 80, 188–207. https://doi.org/10.1006/brln.2001.2588

Myers-Scotton, C. (2006). Natural codeswitching knocks on the laboratory door. *Bilingualism: Language and Cognition*, 9(2), 203–212. https://doi.org/10.1017/S1366728906002549

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge University Press.

Myslín, M., & Levy, R. (2015). Code-switching and predictability of meaning in discourse. *Language*, 91(4), 871–905. https://doi.org/10.1353/lan.2015.0068

Oganian, Y., Korn, C. W., & Heekeren, H. R. (2016). Language switching – but not foreign language use per se – reduces the framing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(1), 140–148. https://doi.org/10.1037/xlm0000161

Opitz, B., & Degner, J. (2012). Emotionality in a second language: it's a matter of time. *Neuropsychologia*, 50(8), 1961–1967. https://doi.org/10.1016/j.neuropsychologia.2012.04.021

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition?. *International Journal of Psychology*, 47(6), 405–428. https://doi.org/10.1080/00207594.2012.743665

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–347. https://doi.org/10.1017/S0140525X12001495

Romero-Rivas, C., Martin, C. D., & Costa, A. (2016). Foreign-accented speech modulates linguistic anticipatory processes. *Neuropsychologia*, 85, 245–255. https://doi.org/10.1016/j.neuropsychologia.2016.03.022

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 8(1), 1–8. https://doi.org/10.1016/S0022-5371(69)80002-2

Ponari, M., Rodríguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing advantage for emotional words in bilingual speakers. *Emotion*, 15(5), 644–652. https://doi.org/10.1037/emo0000061

Poplack, S. (2000). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *The Bilingualism Reader*, 18(2), 221–256. https://doi.org/10.1017/S1366728906002549

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, 61(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raizen, A., Vergis, N., & Christianson, K. (2015). Using eye-tracking to examine the reading of texts containing taboo words. *Interdisciplinary Perspectives on (Im) Politeness*, 213–238. https://doi.org/10.1075/aals.14.11rai

Schacht, A., & Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience*, 9(1), 28–43. https://doi.org/10.3758/CABN.9.1.28

Schotter, E. R., Li, C., & Gollan, T. H. (2019). What reading aloud reveals about speaking: Regressive saccades implicate a failure to monitor, not inattention, in the prevalence of intrusion errors on function words. *Quarterly Journal of Experimental Psychology*, 72(8), 2032–2045. https://doi.org/10.1177/1747021818819480

Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 783. https://doi.org/10.1037/a0027209

Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2014). Emotion words and categories: Evidence from lexical decision. *Cognitive Processing*, 15(2), 209–215. https://doi.org/10.1007/s10339-013-0589-6

Sheikh, N. A., & Titone, D. A. (2013). Sensorimotor and linguistic information attenuate emotional word processing benefits: An eye-movement study. *Emotion*, 13(6), 1107. https://doi.org/10.1037/a0032417

Sheikh, N. A., & Titone, D. (2016). The embodiment of emotional words in a second language: An eye-movement study. *Cognition and Emotion*, 30(3), 488–500. https://doi.org/10.1080/02699931.2015.1018144

Shen, A., Gahl, S., & Johnson, K. (2020). Didn't hear that coming: Effects of withholding phonetic cues to code-switching. *Bilingualism: Language and Cognition*, 1–12. https://doi.org/10.1017/S1366728919000877

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., … & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods Instruments & Computers*, 35(4), 621–633. https://doi.org/10.3758/BF03195542

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., … & Bates, E. (2005). Timed action and object naming. *Cortex*, 41(1), 7–26. https://doi.org/10.1016/S0010-9452(08)70174-6

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … & Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247–250. https://doi.org/10.1016/j.jml.2004.03.002

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Tomić, A. (2015). Use of the English language in English-major students' sociolect: Code switching or code-mixing? In Z. Paunović (ed.), *English language and literature studies: Embracing edges conference proceedings* (pp. 343–357). University of Belgrade.

Tomić, A. & Valdés Kroff, J. R. (2021a). Expecting the unexpected: Code-switching as a facilitatory cue in online sentence processing. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728921000237

Tomić, A. & Valdés Kroff, J. R. (2021b). *Effects of code-switching on emotional processing* [Manuscript in preparation]. Department of Language and Culture, UiT The Arctic University of Norway, Department of Spanish Language and Culture, University of Florida.
University of Michigan English Language Institute. (2006). Michigan English Language Institute College English Test. Retrieved from https://www.michigan-proficiency-exams.com/melicet.html
Travis, R. C. (1936). The latency and velocity of the eye in saccadic movements. In W. R. Miles (ed.), *Psychological monographs: Vol. 47. Psychological studies of human variability* (pp. 242–249). American Psychological Association; Psychological Review Company. https://doi.org/10.1037/13516-018
Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Dussias, P. E. (2018). Experimental contributions of eye-tracking to the understanding of comprehension processes while hearing and reading code-switches. *Linguistic Approaches to Bilingualism*, 8(1), 98–133. https://doi.org/10.1075/lab.16011.val
Van Hell, J. G., Fernandez, C. B., Kootstra, G. J., Litcofsky, K. A., & Ting, C. Y. (2018). Electrophysiological and experimental-behavioral approaches to the study of intra-sentential code-switching. *Linguistic Approaches to Bilingualism*, 8(1), 134–161. https://doi.org/10.1075/lab.16010.van
Zentella, A. C. (1990). Returned migration, language, and identity: Puerto Rican bilinguals in dos worlds/two mundos. *International Journal of the Sociology of Language*, 1990(84), 81–100. https://doi.org/10.1515/ijsl.1990.84.81
Zentella, A. C. (1997). *Growing up bilingual: Puerto Rican children in New York*. Blackwell.

CHAPTER 8

Prediction and grammatical learning in second language sentence processing

Holger Hopp
University of Braunschweig

With a focus on grammatical processing, the chapter surveys recent studies that investigate how second language (L2) learners learn to predict and whether they use prediction for learning the L2 grammar. This chapter first reviews theoretical approaches to the roles of prediction and prediction error as learning mechanisms in first language (L1) and L2 processing. Drawing on priming, (visual-world) eye-tracking and reading-time data, I then discuss how adult L2 learners may differ from monolingual speakers in learning to predict and using prediction for learning. The chapter identifies two key areas in which L2 learning from prediction may be circumscribed, namely, the role of awareness and explicit memory in prediction and the degree to which L2 learners can identify the source of their prediction errors.

Introduction

Research on second language (L2) predictive processing centers on the question of whether late L2 learners recruit anticipatory processing to the same degree as first language (L1) speakers and, if not, which factors may attenuate learners' ability to predict (for overview, Kaan, 2014). A fast growing body of research attests that L2 learners engage fundamentally similar predictive mechanisms as native speakers, albeit they appear to rely less on using grammatical information for prediction. Currently, studies aim to pinpoint the potential causes of reduced grammatical prediction in, e.g., lower automaticity (Grüter et al., 2017), limited cognitive resources (Ito & Pickering, this volume), and more costly and slower lexical processing (Hopp, 2018) of L2 learners versus natives. Yet, much less attention has been paid to the consequences of reduced or different predictive processing among late L2 learners and, specifically, how it may curtail acquisition of the L2 and potentially lead to fossilization. Against the backdrop of models that construe predictive processing as a key learning mechanism, this chapter reviews the links between

https://doi.org/10.1075/bpa.12.08hop

prediction and learning. In focussing on grammatical processing, this chapter defines prediction as the pre-activation of lexical candidates or morphosyntactic, semantic or phonological features of possible lexical candidates, such as syntactic categories, grammatical functions or thematic roles, before a hearer or reader encounters these lexical items in the input.

In the first part of the chapter, I discuss models that link prediction to learning in monolingual acquisition and processing. I then turn to exploring how differences between non-native and native sentence processing may restrict late L2 learners' ability to learn to predict. Finally, I review current processing and priming studies that explore whether and how L2 learners predict to learn.

Grammatical learning and prediction

Most learning models for both child L1 development and (adult) L2 acquisition argue that grammatical learning scopes over a distributional analysis of the input (e.g. Pinker, 1984; Tomasello, 2009; Yang, 2016). In this sense, they are off-line learning theories in that they do not consider the incremental analysis of the input during language comprehension as part of the acquisition process or mechanism.

At the same time, it has long been acknowledged that language processing is critically implicated in learning both as the object of acquisition (Learning to Parse) and as a mechanism of learning (Parsing to Learn; for discussion, see Fodor, 1999). Most accounts construe the role of language processing in acquisition as generating triggers for grammatical acquisition by virtue of parsing failure, i.e. the inability of the current parsing routines to assign a grammatical structure to the incoming input. As a consequence, the parser needs to restructure its processing and its underlying knowledge base to match the format of the input (Omaki & Lidz, 2015).

More recently, different psycholinguistic approaches have contended that predictive processing additionally supports language learning (Christiansen & Chater, 2016; Dell & Chang, 2014; Pickering & Garrod, 2013; Phillips & Ehrenhofer, 2015). In contrast to a parser that passively attempts to integrate the input by assigning a grammatical structure to it, the predictive parser generates active predictions about the input, which will then be confirmed or disconfirmed by the subsequent input. Predictive processing can boost learning in two ways. First, correct predictions during language processing may facilitate learning (a) by confirming and reinforcing knowledge that was predicted and (b) by freeing up resources for attending to and integrating novel information to be learned. For instance, Fernald et al. (2008) report that the efficiency of the predictive processing of adjectives and nouns correlated with learning of novel words that followed the predicted sequence among 2-year-old English-speaking children. In this sense, prediction can act as a helpful "crutch" for learning words and grammar (Pozzan & Trueswell, 2015; Gambi et al., 2020; see also Gambi, this volume). Second, if the learner makes incorrect predictions during language processing, the input can provide the learner with informative feedback on how to restructure their processing and knowledge.

Formal learning approaches construe predictive processing as a way for the learner to test grammatical hypotheses against the input and constrain them to the target grammar after encountering prediction errors (Phillips & Ehrenhofer, 2015). Rather than waiting for phenomena to occur in the input or having to rely on negative evidence or feedback that may not or only inconsistently be available, a predictive learner can actively pitch their predictions against the input and revise their hypotheses to zero in on the target-language grammar. Computationally, such convergence towards the target can be captured in model-based approaches in which choices are made within a set of pre-existing options, such as parameters, that delimit the learner's hypothesis space (e.g. Yang, 2004).

Within connectionist and functional frameworks, predictive processing constitutes the key mechanism by which comprehenders tally their processing preferences to changing input statistics through feedback loops (for discussion, see Myslín & Levy, 2016). To explain learning, these approaches hypothesize that language users make (implicit) predictions about the upcoming input and use prediction error as feedback to restructure their linguistic knowledge and expectations to match the changing statistics in the input so as to minimize future prediction error by way of error-based implicit learning (e.g. Dell & Chang, 2014; Ramscar et al., 2013) or probabilistic belief updating (e.g. Jaeger & Snider, 2013). In these models, the magnitude of prediction error corresponds to the degree of learning, since learning equals the inverse function of the prediction error associated with encountering an unexpected continuation.

In different frameworks, then, the interplay of prediction and the consequences of prediction error has been posited as a central implicit learning mechanism in L1 acquisition and L1 adult processing. At the same time, the evidence that prediction supports learning is largely correlational to date. In child L1 acquisition, children who predict words more efficiently tend to have larger vocabulary sizes (e.g. Borovsky et al., 2012; Mani & Huettig, 2012). Among adult L2 learners, predictive processing sometimes obtains more among highly compared to less proficient L2 learners (e.g. Dussias et al., 2013; Hopp, 2013). However, these correlations leave open whether predictive processing is the cause or the consequence of learning (for discussion, Pickering & Gambi, 2018; Rabagliati et al., 2016).

In a recent study, Reuter et al. (2019) tested 3-to-5-year-old English-speaking children on their processing and learning of labels for novel objects (e.g. *cheem*) that were presented visually alongside familiar objects (e.g. a truck). In visual-world eye-tracking, the participants listened to a semantically constraining context (e.g.

Vroom! You can drive the …) which was followed by the name of the novel object in some of the trials. In a subsequent task measuring the learning of the novel words, the children performed the better the more they had (wrongly) predicted the familiar object and the faster they redirected gazes to the novel object. In conjunction, these findings suggest that the strength of predictions and the subsequent prediction error determined the magnitude of word learning outcomes (see Gambi, this volume, for further discussion).

Although the Reuter et al. study can serve as proof of concept that prediction error can drive learning, the scope of prediction as a learning mechanism may be limited, in particular for the acquisition of grammatical dependencies. Even though basic grammatical prediction of word order and agreement emerges early by age 2 (Gambi et al., 2016), many grammatical dependencies cannot be processed using predictive mechanisms, such as anaphoric relations, in which pronouns refer back to preceding antecedents (e.g. Kaan, 2015). Similarly, for agreement such as gender concord in German, a prenominal determiner only predicts the class of the following noun (e.g. *der*$_{\text{MASC}}$ → N$_{\text{MASC}}$), whereas there is a high backward transitional probability between a particular noun and a determiner marking its gender (e.g. *Knochen*, 'bone'$_{\text{MASC}}$ → *der*$_{\text{MASC}}$). Such grammatical dependencies and their constraints are arguably better learnt from a distributional analysis of syntactic relations in the input. Moreover, children may not have the resources to quickly make complex predictions, especially those that involve the integration of multiple types of information. For complex long-distance filler-gap dependencies, children, unlike adults, often do not show predictive processing, even though they have adult-like knowledge of these dependencies (e.g. Atkinson et al., 2018). In other studies, children take much longer than adults to integrate different information in the predictive processing of cross-clausal contingencies (e.g. Hartshorne et al., 2015). Given that children have knowledge of these dependencies but do not use it for prediction, it is not very likely that this knowledge had initially been acquired using predictive processing. Instead, prediction may be the outcome of learning. Finally, even when they make predictions, it is an open question how much children could learn from a prediction error, seeing that even 7-to-10-year-old children struggle to revise their initial processing (e.g. Trueswell et al., 1999; see Phillips & Ehrenhofer, 2015, for review). As children's revision abilities develop late, the impact of learning mechanisms driven by prediction error may be limited, at least in the first years of life.

Learning and prediction in L2 acquisition

Among (adult) L2 learners, the conditions for prediction are different, since L2 learners are cognitively more mature and the basic mechanisms of predictive processing have been acquired during L1 acquisition and are fundamentally in place in L2 acquisition and processing (Kaan, 2014).

However, grammatical development in the L2 via prediction may be affected by two differences between L1 and L2 predictive processing that tie in with the dual nature of prediction as an object and a mechanism of learning, respectively.

– L2 users may not learn to make native-like grammatical predictions in the L2 (Learning to Predict)
– L2 users may not learn from predictions in the L2 like native speakers (Predicting to Learn)

In the following, I review research on how L2 learners learn to predict and how they can recruit prediction as a learning mechanism in the acquisition of L2 grammar.

Learning to predict in an L2

Not all predictions need to be learned by L2 learners. Lexical-semantic prediction emerges early in L2 acquisition, presumably due to L1 transfer, as suggested by L1 effects on semantic prediction (Van Bergen & Flecken, 2017). When it comes to grammatical prediction, though, unless learners transfer analogous grammatical properties from the L1, they need to acquire the relevant grammatical knowledge in the L2.

This is trivial in the sense that if a learner does not know, e.g. the form or function of a classifier in Chinese, they cannot use it predictively. Yet, if L2 learners categorize lexical classes or grammatical functions from the input differently than native speakers, they will also learn to use grammatical prediction differently to native speakers. For instance, English learners of Chinese establish more coarse-grained classifier systems and make predictions in accordance with their subjective grammatical knowledge (e.g. Grüter et al., 2020, see also Lemhöfer et al., 2014).

Learning grammatical prediction may also be affected by misaligned L1 properties (e.g. Dussias et al., 2003). For L1 Russian child and adult learners of German, Hopp and Lemmerth (2018) and Lemmerth and Hopp (2019) reported that lexical gender congruency between German and Russian affected predictive gender processing in that learners used German gender on articles or adjectives predictively for nouns that overlapped in gender class between L1 and L2 (e.g.

Vase$_{\text{FEM}}$ 'vase' – Russian: *ва́за*$_{\text{FEM}}$), yet not those that differed in gender between L1 and L2 (e.g. *Knochen*$_{\text{MASC}}$ 'bone' – *кость*$_{\text{FEM}}$). As suggested by current models of the bilingual mental lexicon (e.g. Dijkstra & van Heuven, 2002), activation of the gender of an article or adjective spreads non-selectively across languages to lexical candidates across both languages and, e.g., a masculine article will pre-activate not only the German masculine word for 'bone' but also Russian word candidates with masculine gender, while simultaneously suppressing activation of 'bone', since its gender is not masculine in Russian (see also Morales et al., 2016). Hence, interfering knowledge of the L1 may initially affect and diffuse grammatical prediction among L2 learners and delay its acquisition (see also Foucart, this volume).

Yet, learning to predict extends beyond knowledge or classification of the specific form, feature or lexical items involved in prediction. In a series of studies, Hopp (2013, 2016), investigated variation in the use of grammatical gender for prediction from German articles to nouns (*der*$_{\text{MASC}}$ → *kleine*$_{\text{ambiguous}}$ *Knochen*$_{\text{MASC}}$) and adjectives to nouns (*ein*$_{\text{ambiguous}}$ *klein-er*$_{\text{MASC}}$ → *Knochen*$_{\text{MASC}}$; 'the/a small bone'). For L1 English learners, the studies found that only L2 learners who had target knowledge of the lexical genders of all nouns in the experiment (i.e. *Knochen* = MASC, *Vase* = FEM, etc.) could use gender for prediction, while learners who wavered between gender options for some nouns (e.g. *Knochen* = MASC or NEUT) did not demonstrate gender prediction even for the majority of nouns for which they had robust gender knowledge (e.g. *Vase* = FEM). Following accounts that prediction is driven by utility and affected by the reliability of the input (e.g. Brothers et al., 2019), the L2 studies on gender suggest that variable L2 knowledge impairs the usefulness of prediction, since it leads to frequent mispredictions that require costly reanalysis (for discussion, Kuperberg & Jaeger, 2016). In consequence, even partially lacking, inconsistent or non-target knowledge may reduce or cancel predictive processing (see also Hanulikova et al., 2012). Hence, the learning of prediction may be held back by the low success of learners' initial grammatical predictions.

Finally, a learner needs to have the resources to make grammatical predictions. For instance, working memory capacity affects the speed and time course of grammatical prediction (e.g. Huettig & Janse, 2016), and processing an L2 is inherently more taxing than comprehending the L1. In addition, L2 learners are compromised in predictive processing under heightened task demands (Ito et al., 2018) or at faster speech rates (for discussion, see Huettig & Guerra, 2019; Ito & Pickering, this volume). Hence, the greater strains of predictive processing on cognitive resources may lead to predictive processing emerging relatively late in the course of L2 acquisition. Relatedly, resources may foremost be allocated to or exhausted by the bottom-up integration of information, so that (top-down) prediction does not emerge as quickly or fully in L2 learners as among child L1 learners or adult monolinguals (e.g. Grüter et al., 2017; see also Phillips & Ehrenhofer, 2015).

In sum, learners do not necessarily learn to predict in a native-like way, even though native-like grammatical prediction may be achieved (Hopp, 2013; Hopp & Lemmerth, 2018; Schlenter & Felser, this volume). L2 learners may make no or slow grammatical predictions due to lacking or non-target L2 knowledge, the unreliability of their predictions or processing limitations. Moreover, they may generate partially erroneous predictions based on non-target analyses of the L2 input or the transfer of L1 properties.

Learning to predict due to exposure and structural priming

Against this backdrop, several studies that involved the presentation of massed input in natural and artificial languages examined whether L2 learners learn to make grammatical predictions from exposure. In a pre-posttest-training study on the gender in article-noun combinations, intermediate L1 English learners demonstrated predictive gender processing in German after massed exposure to the target article noun combinations, suggesting that input exemplifying local co-occurrence regularities between articles and nouns triggers predictive processing of agreement relations in L2 development (Hopp, 2016). In a similar vein, Curcic et al. (2019) report that prediction from article-like forms co-occuring with sets of nouns emerges after some exposure in a miniature artificial language. These findings from visual world eye-tracking are consistent with ERP studies on the processing of gender-like agreement in artificial languages (for review, see Morgan-Short, 2020). Yet, the generalizability of these findings is somewhat circumscribed in that successful gender prediction also implicates (explicit) lexical knowledge of the gender classes of nouns and, in artificial language, is confounded with word learning.

Beyond grammatical gender, Hopp (2020: Experiment 2) studied the effects of exposure on the processing of tense mismatches between lexical tense marking on adverbials and grammatical tense marking on verbs (1), following Roberts and Liszka (2013). In a self-paced reading task, the question was whether L2 learners develop predictions for grammatical tense marking on the basis of frequent exposure to matching temporal adverbials.

(1) a. Last week/*since the summer, James went swimming every day. (Past tense)
b. *Last week/since the summer, James has gone swimming every day. (Present perfect)

In a pretest, intermediate-level L1 German learners of English were not sensitive to tense mismatches in self-paced reading. Following an exposure phase that exemplified the target pairings between adverbials and grammatical tense, the experimental group demonstrated reading slowdowns for tense mismatches on past tense verbs in the posttest, which suggests that learners had learned to predict the verb tense from the sentence-initial temporal adverbial. In contrast, the control group, which read lexically matched sentences that did not illustrate variation in tenses, remained insensitive to tense mismatches throughout. Of note, Hopp (2020) administered off-line cloze and acceptability judgment tasks to both groups after the posttest. In neither task did the experimental group differ from the control group, which indicates that the development of predictive processing did not translate into group differences in more explicit knowledge tasks testing awareness of tense mismatches.

Hence, the relations between real-time prediction and the development of explicit grammatical knowledge or awareness are not straightforward. As a case in point, Andringa (2020) tested L1 Dutch learners of an artificial miniature language that used determiners as animacy and distal markers. In a visual-world task, the participants were presented with images of (in)animate objects in close or far proximity to the viewer in seven blocks. In a debriefing after the experiment, the experimenter assessed in which block the participant had become aware of the target structure. Subsequently, the preceding blocks were collectively analysed as "unaware" blocks and the following blocks as "aware" blocks. The results showed that only when learners had become aware of the target structure did they use determiners predictively. Moreover, participants who remained unaware throughout did not manifest any predictive use of determiner information (see also Curcic et al., 2019). In ERP experiments, L1 adults also demonstrated enhanced lexical prediction when participants were explicitly invited to engage in prediction (e.g. Brothers et al., 2017), suggesting that explicit strategies generally enhance prediction.

Further evidence of the involvement of explicit strategies in prediction comes from structural priming studies, for instance, a recent study by Jackson and Hopp (2020), discussed below. Structural priming refers to the tendency of comprehenders or speakers to reuse a recently experienced syntactic structure (for review, Jackson, 2018; Van Gompel & Arai, 2018). For instance, if someone hears descriptions involving ditransitive verbs being used with a prepositional object (e.g. *Mary gave the ice cream to the boy*), they are more likely to reuse this construction than a double-object continuation (*Mary gave the boy the ice cream*; Bock, 1986; for review, Mahowald et al., 2016). Critically, less frequent structures, e.g. passives, show stronger priming than more frequent structures, e.g. active sentences. In models of priming as error-driven implicit learning (e.g. Dell & Chang, 2014), this so-called inverse frequency effect of priming has been taken to suggest that prediction and prediction error underlie priming, since the magnitude of priming for a structure is inversely correlated with the strength of its prediction, as operationalized in its frequency.

When applied to L2 acquisition, the implicit-learning account of structural priming leads to the expectation that the magnitudes of priming differ across languages and between L1 and L2. Building on this logic, Jackson and Hopp (2020) used across-language comparisons in structural priming to investigate the relation between the size of the prediction error and implicit learning via priming. They compared the short-term and long-term priming of fronted adverbials among native English speakers as well as L1 German learners of English in German and English (*In the morning the grandfather drinks hot chocolate*). As per the inverse frequency effects for priming, they hypothesized that encountering a lower frequency construction will give rise to a larger prediction error. Since fronted temporal adverbials are less frequent in English than in German, English natives should show greater priming than German natives, and L2 speakers of English should demonstrate the largest priming, since they encounter fronted adverbials even less often in English than natives. The short-term priming magnitudes patterned along with these predictions, consistent with the assumption that the size of the prediction error determines the extent of priming. However, in longer-term priming, i.e. the differences in the (unprimed) production of fronted adverbials between a pretest and an immediate posttest following the priming phase, only the results for the L1 groups mirrored the magnitude differences in short-term priming (English > German), whereas the L2 group had the smallest priming effect. In other words, for the L2 group, the size of the prediction error experienced in short-term priming did not translate into corresponding magnitudes in longer-term priming.

These findings are in line with two-locus accounts of priming (e.g. Ferreira & Bock, 2006), according to which short-term priming reflects explicit memory processes, while long-term priming represents implicit learning. Under such an account, short-term priming and implicit learning would rely on different mechanisms among L2 learners and, as a consequence, prediction error would not feed directly into implicit learning in an L2.[1] Such an account is consistent with findings from lower-proficiency L2 learners demonstrating short-term structural priming even though they do not have abstract representations of the primed structure in long-term memory (Bernolet et al., 2013; Jackson & Ruf, 2017). Moreover, these discontinuities between short-term and long-term priming resemble the lack of across-task transfer of the adaptation effects among L2 learners in Hopp (2020: Experiment 2), and explicit memory strategies can also account for the learning of gender prediction in studies on natural (Hopp, 2016) and artificial (Curcic et al., 2019) languages. Finally, in studies on explicit sentence priming, L2 learners who are asked to make strategic memory-based predictions about the structure of a prime sentence demonstrate larger priming effects than participants who only passively repeat prime sentences (Grüter, Zhu, & Jackson, this volume).

1. Note that L1-L2 differences do not entail that L2 learners are restricted to one memory system, since both systems are clearly available, as demonstrated in cumulative priming (e.g. Kaan & Chun, 2018a) and long-term priming (e.g. Shin & Christianson, 2012) among L2 learners.

Clearly, more research into variation of prediction across tasks, between L2 and L1 speakers as well as between L2 learners is necessary. Importantly, moving research forward from asking *whether* L2 learners learn to predict to *how* they learn to predict can also relate the role of prediction in learning to central debates in second language acquisition about the recruitment of explicit versus implicit knowledge (e.g. Rebuschat, 2015), declarative versus procedural memory (Morgan-Short & Ullman, to appear) as well as incidental vs. intentional learning in statistical learning (Rebuschat & Williams, to appear).

Predicting to learn in an L2

Lower degrees or different types of prediction among L2 learners also have consequences for the use of prediction as a learning mechanism. For one thing, lesser or slower prediction means that L2 learners cannot harness feedback since they do not generate any prediction error in the first place that could lead them to revise their hypotheses about the target-language grammar. For another, different predictions may set them up to experience different prediction errors and miss the chance to revise their processing towards the target. Many studies on L2 sentence processing report that late L2 learners overrely on non-grammatical information in sentence processing, often resorting to grammatically "shallow" processing (Clahsen & Felser, 2006, 2018), which means that they readily make lexical-semantic predictions but underrely on grammatical information. With a view to L2 grammatical development, Dekydtspotter and Renaud (2014: 145) argue that "[…] whenever L2 sentence processing is shallow (i.e. not implicating the generation of detailed representations), L2 learners will not have access to the learning triggers that reside in those details." When applied to predictive processing, shallow processing entails that a prediction error generated by semantic predictions may be of a different type, magnitude or consequence than a prediction error occasioned by a grammar-based prediction.

Finally, when learners experience a prediction error, they need to revise the prediction as part of error-based implicit learning. For revision, however, L2 learners have been argued to suffer greater difficulty in abandoning and revising an initial parse, compared to L1 adults (e.g. Jacob & Felser, 2016; Safak & Hopp, 2021). In the context of predictive processing, such problems with reanalysis may entail that L2 learners persist with the initial misanalysis (Pozzan & Trueswell, 2016) and fail to revise their predictions even after multiple encounters of prediction errors (e.g. Kaan et al., 2019).

To illustrate these points, a study by Hopp (2015) may serve as a useful example. In this study, L1 English learners of L2 German were tested on the interactive use of semantic and grammatical prediction using sentences as in (2).

(2) a. Der Wolf tötet gleich den Hirsch.
The$_{NOM}$ wolf kills soon the$_{ACC}$ deer
'The wolf soon kills the deer'
b. Den Wolf tötet gleich der Jäger.
The$_{ACC}$ wolf kills soon the$_{NOM}$ hunter
'The hunter soon kills the wolf'

As speakers of a free word order language, native Germans reliably used case marking (nominative vs accusative) on the first NP designating its syntactic function as subject or object in anticipating a patient (object) as the postverbal noun in (2a), e.g. the deer, and predicting an agent (subject) noun in (2b), e.g. the hunter, when they heard the first NP and the verb. Hence, they integrated morphosyntactic case information on articles and semantic event information encoded in the verb for prediction (see also Kamide, Scheepers, & Altmann, 2003). In contrast, the L2 learners across proficiency levels predicted the second NP to be the patient (object) noun based on the semantics of the verb, irrespective of whether the first NP bore nominative or accusative case marking.

First, this study bears out asymmetries in the information types used for prediction among L2 learners in that L2 learners made early and robust use of semantic information encoded in the verb (see also Dijkgraaf et al., 2017; Ito et al., 2018); yet, they did not use case marking for prediction (see also Mitsugi & MacWhinney, 2016; but, see Schlenter & Felser, this volume, and Frenck-Mestre et al., 2019, for L1 effects). Their semantic predictions set them up for frequent prediction errors in sentences like (2b). However, the L2 learners were slower to revise the semantically based prediction than natives even when the disambiguating lexical information of the second NP became available, suggesting that they had trouble revising their predictions. Third, the fact that prediction towards the patient was as strong in the OVS as in the SVO conditions suggests that participants did not abandon making erroneous predictions for the second NP to be the patient and revise their predictions, despite encountering frequent prediction errors in the course of the experiment.

The lack of using case marking for prediction among the L2 learners did not seem related to lack of knowledge of case markers (see also Mitsugi & MacWhinney, 2016) or the low salience of case marked articles. In a follow-up study, Henry et al. (2020) report that the difficulty in using case marking persists even when the salience of case marking was enhanced through prosodic cues. Instead, L2 learners may not have parsed the case marking as per shallow processing in the first place, they may not have been able to use the prediction error as a signal telling them how to revise their parses or they may not have had the computational means to revise their predictions in real time. Each and all of these factors may constrain their chances to learn from prediction error.

Syntactic adaptation and the consequences of prediction error

The roles of prediction and prediction error in L2 learning are beginning to be addressed in studies on syntactic adaptation. Syntactic adaptation refers to processing adjustments in response to input changes in the environment, and studies on syntactic adaptation typically involve temporarily ambiguous sentences that have a more frequent and thus preferred variant (for review, Kaan & Chun, 2018b). Experiments manipulate the input statistics to investigate whether massed exposure to the dispreferred, i.e. non-predicted, variant eases processing difficulty relative to the preferred variant over time.

In several reading experiments on syntactic ambiguities, native speakers were found to adapt in main clause versus reduced relative clause ambiguities (Fine et al., 2013; Yan & Jaeger, 2020; though see Harrington Stack et al., 2018), coordination ambiguities or filler-gap dependencies (Kaan et al., 2019). For instance, Kaan et al. (2019) found that native English readers adapted their processing after having encountered multiple instances of sentences involving clausal coordination as in (3).

(3) The servant cleaned the table and the floor was cleaned by the maid.

Compared to an unambiguous type of clausal coordination with the conjunction *but*, the differences in reading times at the verb of the second clause (*was cleaned*) decreased over the course of the experiment. In contrast, a high-proficiency L2 group did not show analogous adaptation to the dispreferred clausal coordination. To account for the lack of adaptation in the L2 group, Kaan et al. argued that the signal from prediction error may have been noisier and thus less effective in engendering adaptation in the L2 group versus the natives.[2]

Hopp (2020) examined how exposure to input that was designed to induce prediction error *and* to guide learners in the right direction in the processing of temporary object-subject ambiguities (4) can change predictions during L2 processing.

(4) a. When the girl was playing, the boy made some funny noises. (Control)
b. When the girl was playing he made some funny noises. (Case)
c. When the girl was playing the boy made some funny noises. (Implausible)
d. When the girl was playing the piano made some funny noises. (Plausible)

Using a pre-post-test control-group design, the study built on earlier findings in Hopp (2014) that L1 German learners were not sensitive to a nominative case-marked pronoun (*he*) signalling the intransitive use of the preceding verb in

2. Note that Kaan et al. (2019) also tested filled gap constructions for which they did not find any adaptation effects in the native speaker group either. Hence, task or construction-specific effects appear to have affected adaptation on top of L1-L2 differences.

eye-tracking during reading. The learners in Hopp (2020) equally showed longer reading times on the main clause verb *made* for (4b) than (4a), suggesting that they temporarily analysed the pronoun as the object of the verb *playing*, which they then needed to reanalyse towards the subject when encountering *made*. Subsequently, the experimental group was exposed to sentences exemplifying both the intransitive use of the first verb and the use of the nominative pronoun as the subject (*The boy played and he pleased the parents with the music.*). The control group read lexically-matched sentences in which the verb was used transitively and the pronoun was ambiguous between nominative and accusative case (*The boy played the music and it pleased the parents.*). In an immediate posttest, the experimental group displayed a reduction in reading times on the main clause verb in (4b), paralleling the reading times of the control condition (4a), while the control group continued to get garden-pathed with sentences (4b). Critically, the intransitive uses of the verbs in the exposure phase did not lead to a blanket reduction in reanalysis costs, since the participants in the experimental group continued to garden-path in (4c) and (4d). Hence, simple exposure to intransitive uses of the verb did not cancel the prediction that an optionally transitive verb be followed by an object. In line with accounts that the nature of the parsing error and the difficulty of their repair affects processing revisions (Fodor & Inoue, 1994), Hopp argued that the combination of the intransitive use of the verb and the unambiguous flagging of the pronoun as a grammatical subject underlay adaptation. In other words, only when the prediction error came with its solution did L2 learners change their grammatical predictions towards integrating case marking.

This interpretation squares with the findings in the Hopp (2015) study on case marking discussed above where prediction error was not accompanied by any information that pointed the learners to the source of the processing problem and its solution; as a consequence, no adaptation and learning occurred. In contrast, other studies that did find adaptation effects with syntactic ambiguities in L2 learners comprised exclusively structural phenomena, such as reduced vs. non-reduced relative clauses (Arai, 2016), relative clause attachment ambiguities (Chun, 2020; Chun et al., this volume), or structures disambiguated by semantic, e.g. animacy, information in object versus subject relative clauses (Nitschke et al., 2014). Learners may more easily reanalyse such parses than identifying and rectifying predictions that require the integration of different types of, e.g. lexical, syntactic, and morphological information.

In order to learn from prediction error, then, learners critically need to be able to track the grammatical source of the error, such that they can revise their predictions to zero in on the target grammar. Otherwise, the learner will be at a loss as what to learn, and the prediction error will not allow for grammatical input to filter into learner intake for L2 development. Hence, unrevised prediction errors may lead to persistent misparsing and fossilization in the L2. The emerging findings from studies on L2 syntactic adaptation suggest that, rather than merely studying *whether* L2 learners can adapt via prediction error, probing *how* they learn from prediction error may uncover insights into the limits of L2 learning.

Conclusions and outlook

In this chapter, I reviewed links between prediction and the learning of L2 grammar. On the one hand, L2 learners need to learn to predict grammatically, which may be constrained by lacking or inconsistent lexical and grammatical knowledge, L1 influence, shallow parsing and cognitive resource limitations. On the other hand, L2 learners can only learn from prediction error if they make a prediction in the first place, if they can identify the source of the prediction error, if they can revise their grammatical processing accordingly and if prediction error feeds into long-term memory supporting implicit learning.

In order to better understand the role of prediction in learning and the relative impacts of the above factors, we require longitudinal studies on the emergence and development of predictive processing, and particularly studies that directly link prediction error to grammatical learning at the initial stages of L2 acquisition.

At lower proficiency levels, the learner has little experience with the L2 so that unexpected input should be associated with a strong prediction error, which, in turn, should occasion a large degree of learning. At the same time, lower-proficiency learners may lack the grammatical knowledge to generate and the processing skills to rapidly implement predictions in processing the L2. Conversely, more advanced learners can make more predictions and can more easily make predictions to learn from, yet they have less to learn in the L2. Hence, mapping the development of learning by prediction will delineate the scope of predictive processing as a learning mechanism in L2 acquisition.

Finally, the field should move towards across-population comparisons in predictive processing. Obvious comparisons involve early vs. late (L2) learners and studying different L1 groups within adult L2 learning. Others comprise (L2) learners and speakers with developmental language disorders (e.g. Jones & Westermann, 2020) in order to assess the role of predictive processing as a learning mechanism in another population that does not invariably reach high grammatical proficiency. Studies along these lines will not only elucidate the links between prediction and learning but also contribute to identifying differences or specifics in L2 predictive processing.

References

Andringa, S. (2020). The emergence of awareness in uninstructed L2 learning: A visual world eye tracking study. *Second Language Research*. https://doi.org/10.1177/0267658320915502

Arai, M. (2016). Importance of making (prediction) errors in L2 processing. Poster presented at the 22nd Annual Conference on Architectures and Mechanisms for Language Processing, Bilbao, Spain, September 2016.

Atkinson, E., Wagers, M. W., Lidz, J., Phillips, C., & Omaki, A. (2018). Developing incrementality in filler-gap dependency processing. *Cognition*, 179, 132–149. https://doi.org/10.1016/j.cognition.2018.05.022

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, 112(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Brothers, T., Dave, S., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2019). Flexible predictions during listening comprehension: Speaker reliability affects anticipatory processes. *Neuropsychologia*, 135, 107225. https://doi.org/10.1016/j.neuropsychologia.2019.107225

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, 93, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39, E62. https://doi.org/10.1017/S0140525X1500031X

Chun, E. (2020). L2 prediction guided by linguistic experience. *English Teaching*, 75, 79–103. https://doi.org/10.15858/engtea.75.s1.202006.79

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42. https://doi.org/10.1017/S0142716406060024

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, 40(3), 693–706. https://doi.org/10.1017/S0272263117000250

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, 69, 42–71. https://doi.org/10.1111/lang.12321

Dekydtspotter, L., & Renaud, C. (2014). On second language processing and grammatical development: The parser in second language acquisition. *Linguistic Approaches to Bilingualism*, 4(2), 131–165. https://doi.org/10.1075/lab.4.2.01dek

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120394. https://doi.org/10.1098/rstb.2012.0394

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dijkstra, A. F. J., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197. https://doi.org/10.1017/S1366728902003012

Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). John Benjamins. https://doi.org/10.1075/lald.44.06fer

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*, 21(7–8), 1011–1029. https://doi.org/10.1080/01690960600824609

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PloS one*, 8(10), e77661. https://doi.org/10.1371/journal.pone.0077661

Fodor, J. D. (1999). Learnability theory: Triggers for parsing with. In E. C. Klein & G. Martohardjono (Eds.), *The development of second language grammars: A generative approach* (p. 373–406). John Benjamins. https://doi.org/10.1075/lald.18.17fod

Fodor, J. D., & Inoue, A. (1994). The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research*, 23(5), 407–434. https://doi.org/10.1007/BF02143947

Frenck-Mestre, C., Kim, S. K., Choo, H., Ghio, A., Herschensohn, J., & Koh, S. (2019). Look and listen! The online processing of Korean case by native and non-native speakers. *Language, Cognition and Neuroscience*, 34(3), 385–404. https://doi.org/10.1080/23273798.2018.1549332

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in preschoolers' linguistic predictions. *Cognition*, 157, 340–351. https://doi.org/10.1016/j.cognition.2016.10.003

Gambi, C., Pickering, M. J., & Rabagliati, H. (2020). The relation between preschoolers' vocabulary development and their ability to predict and recognize words. *Child Development*. https://doi.org/10.1111/cdev.13465

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, 7(2), 199–229. https://doi.org/10.1075/lab.15011.gru

Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in L1 and L2 Chinese: The role of semantic and grammatical cues. *Language, Cognition and Neuroscience*, 35(2), 221–234. https://doi.org/10.1080/23273798.2019.1648840

Hanulíková, A., Van Alphen, P. M., Van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, 24(4), 878–887. https://doi.org/10.1162/jocn_a_00103

Harrington Stack, C. M., James, A. N., & Watson, D. G. (2018). A failure to replicate rapid syntactic adaptation in comprehension. *Memory & Cognition*, 46(6), 864–877. https://doi.org/10.3758/s13421-018-0808-6

Hartshorne, J. K., Nappa, R., & Snedeker, J. (2015). Development of the first-mention bias. *Journal of Child Language*, 42(2), 423–446. https://doi.org/10.1017/S0305000914000075

Henry, N., Hopp, H., & Jackson, C. N. (2020). Cue coalitions and additivity in predictive processing: the interaction between case and prosody in L2 German. *Second Language Research*. https://doi.org/10.1177/0267658320963151

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2014). Individual differences in the L2 processing of object-subject ambiguities. *Applied Psycholinguistics*, 36(2), 129–173. https://doi.org/10.1017/S0142716413000180

Hopp, H. (2015). Semantics and morphosyntax in L2 predictive sentence processing. *International Review of Applied Linguistics*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014
Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in adult L2 acquisition. *Second Language Research*, 32(2), 277–307. https://doi.org/10.1177/0267658315624960
Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*, 17, 5–27. https://doi.org/10.11431/secondlanguage.17.0_5
Hopp, H. (2020). Morphosyntactic adaptation in adult L2 processing: Exposure and the processing of case and tense violations. *Applied Psycholinguistics*, 41(3), 627–656. https://doi.org/10.1017/S0142716420000119
Hopp, H., & Lemmerth, N. (2018). Lexical and syntactic congruency in L2 predictive gender processing. *Studies in Second Language Acquisition*, 40(1), 171–199. https://doi.org/10.1017/S0272263116000437
Huettig, F., & Guerra, E. (2019). Effects of speech rate, preview time of visual context, and participant instructions reveal strong limits on prediction in language processing. *Brain Research*, 1706, 196–208. https://doi.org/10.1016/j.brainres.2018.11.013
Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, 31(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459
Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050
Jackson, C. N. (2018). Second language structural priming: A critical review and directions for future research. *Second Language Research*, 34(4), 539–552. https://doi.org/10.1177/0267658317746207
Jackson, C. N., & Hopp, H. (2020). Prediction error and implicit learning in L1 and L2 syntactic priming. *International Journal of Bilingualism*, 24(5–6), 895–911. https://doi.org/10.1177/1367006920902851
Jackson, C. N., & Ruf, H. T. (2017). The priming of word order in second language German. *Applied Psycholinguistics*, 38(2), 315–345. https://doi.org/10.1017/S0142716416000205
Jacob, G., & Felser, C. (2016). Reanalysis and semantic persistence in native and non-native garden-path recovery. *Quarterly Journal of Experimental Psychology*, 69(5), 907–925. https://doi.org/10.1080/17470218.2014.984231
Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013
Jones, S. D., & Westermann, G. (2020, May 15). Predictive processing and developmental language disorder. https://doi.org/10.31234/osf.io/gzk9n
Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa
Kaan, E. (2015). Knowing without predicting, predicting without learning. *Linguistic Approaches to Bilingualism*, 5(4), 482–486. https://doi.org/10.1075/lab.5.4.07kaa
Kaan, E., & Chun, E. (2018a). Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, 21(2), 228–242. https://doi.org/10.1017/S1366728916001231
Kaan, E., & Chun, E. (2018b). Syntactic adaptation. In D. Watson & K. Federmeier (Eds.), *Psychology of learning and motivation* (Vol. 68, pp. 85–116). Academic Press.

Kaan, E., Futch, C., Fuertes, R. F., Mujcinovic, S., & de la Fuente, E. Á. (2019). Adaptation to syntactic structures in native and nonnative sentence comprehension. *Applied Psycholinguistics*, 40(1), 3–27. https://doi.org/10.1017/S0142716418000437

Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32(1), 37–55. https://doi.org/10.1023/A:1021933015362

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Lemhöfer, K., Schriefers, H., & Indefrey, P. (2014). Idiosyncratic grammars: Syntactic processing in second language comprehension uses subjective feature representations. *Journal of Cognitive Neuroscience*, 26(7), 1428–1444. https://doi.org/10.1162/jocn_a_00609

Lemmerth, N. & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Effects of lexical and syntactic cross-linguistic influence. *Language Acquisition*, 26(1), 21–45. https://doi.org/10.1080/10489223.2017.1391815

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, 91, 5–27. https://doi.org/10.1016/j.jml.2016.03.009

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843–847. https://doi.org/10.1037/a0029284

Mitsugi, S., & Macwhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881

Morales, L., Paolieri, D., Dussias, P. E., & Valdés-Kroff, J. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition*, 19(2), 294–310. https://doi.org/10.1017/S1366728915000176

Morgan-Short, K. (2020). Insights into the neural mechanisms of becoming bilingual: A brief synthesis of second language research with artificial linguistic systems. *Bilingualism: Language and Cognition*, 23(1), 87–91. https://doi.org/10.1017/S1366728919000701

Morgan-Short, K. & Ullman, M. T. (to appear). Declarative and procedural memory in second language learning: Psycholinguistic considerations. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics*. London: Routledge.

Myslín, M., & Levy, R. (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, 147, 29–56. https://doi.org/10.1016/j.cognition.2015.10.021

Nitschke, S., Serratrice, L., & Kidd, E. (2014). The effect of linguistic nativeness on structural priming in comprehension. *Language, Cognition and Neuroscience*, 29(5), 525–542. https://doi.org/10.1080/01690965.2013.766355

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, 22(2), 158–192. https://doi.org/10.1080/10489223.2014.943903

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, 5(4), 409–453. https://doi.org/10.1075/lab.5.4.01phi

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(04), 329–347. https://doi.org/10.1017/S0140525X12001495

Pinker, S. (1984). *Language learnability and language development*. Harvard University Press.

Pozzan, L., & Trueswell, J. C. (2015). Syntactic processing and acquisition. *Linguistic Approaches to Bilingualism*, 5(4), 516–521. https://doi.org/10.1075/lab.5.4.13poz

Pozzan, L., & Trueswell, J. C. (2016). Second language processing and revision of garden-path sentences: A visual word study. *Bilingualism: Language and Cognition*, 19(3), 636–643. https://doi.org/10.1017/S1366728915000838

Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience*, 31(1), 94–105. https://doi.org/10.1080/23273798.2015.1077979

Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of" mouses" in adult speech. *Language*, 760–793. https://doi.org/10.1353/lan.2013.0068

Rebuschat, P. (Ed.). (2015). *Implicit and explicit learning of languages*. John Benjamins. https://doi.org/10.1075/sibil.48

Rebuschat, P. & Williams, J. (to appear). The psycholinguistics of L2 learning: A cognitive psychology perspective. In A. Godfroid & H. Hopp (Eds.). *The Routledge handbook of second language acquisition and psycholinguistics*. London: Routledge.

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*, 55(8), 1656–1665. https://doi.org/10.1037/dev0000754

Roberts, L., & Liszka, S. A. (2013). Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, 29(4), 413–439. https://doi.org/10.1177/0267658313503171

Safak, D. F. & Hopp, H. (2021). Verb bias and semantic persistence effects in L2 ambiguity resolution. *Second Language Research*. https://doi.org/10.1177/0267658321997904

Shin, J. A., & Christianson, K. (2012). Structural priming and second language learning. *Language Learning*, 62(3), 931–964. https://doi.org/10.1111/j.1467-9922.2011.00657.x

Tomasello, M. (2009). *Constructing a language*. Harvard University Press.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032-3

Van Bergen, G., & Flecken, M. (2017). Putting things in new places: Linguistic experience modulates the predictive power of placement verb semantics. *Journal of Memory and Language*, 92, 26–42. https://doi.org/10.1016/j.jml.2016.05.003

Van Gompel, R. P., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, 21(3), 448–455. https://doi.org/10.1017/S1366728917000542

Yan, S., & Jaeger, T. F. (2020). Expectation adaptation during natural reading. *Language, Cognition and Neuroscience*, 1–29. https://doi.org/10.1080/23273798.2020.1784447

Yang, C. D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*, 8(10), 451–456. https://doi.org/10.1016/j.tics.2004.08.006

Yang, C. (2016). *The price of linguistic productivity*. The MIT Press. https://doi.org/10.7551/mitpress/9780262035323.001.0001

CHAPTER 9

The role of prediction in second language vocabulary learning

Chiara Gambi
Cardiff University

While we have a good understanding of how prediction is implicated in processing utterances in a second language, the idea that prediction may also be implicated in learning a second language has so far received very little attention. I briefly review the evidence for and against prediction as a fundamental mechanism of language acquisition in childhood, focusing in particular on the links between children's prediction skills and their vocabulary knowledge. I then illustrate how prediction, and in particular prediction errors, may drive learning of new words in adults, by supporting the encoding and consolidation of representations for these words in memory. Finally, I discuss how prediction-based mechanisms might be implicated in generating and sustaining motivation to learn a second language.

Overview

Prediction is now seen as an important mechanism in first and second language processing (Huettig, 2015; Kaan, 2014; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Pickering & Garrod, 2013). For the purpose of this chapter, I define prediction as the pre-activation of linguistic representations ahead of bottom-up processing of the linguistic input at the corresponding representational level (e.g., syntax, phonology; Pickering & Gambi, 2018). In this chapter, I will mainly focus on the downstream consequences of this pre-activation, which are measurable once the comprehender processes input that either matches or mismatches the pre-activated representations. How do matching or mismatching predictions affect the long-term storage of linguistic information in memory? I will ask this question to understand what role, if any, prediction plays in the learning of a second language (L2), and specifically in consolidating L2 vocabulary in memory (see also Hopp, this volume). Because evidence surrounding the role of prediction in L2 learning is still limited, I will begin with an overview of what we know about the role of prediction during

https://doi.org/10.1075/bpa.12.09gam

first language (L1) vocabulary development in childhood. I will then review the (still-limited) evidence for the role of prediction in encoding and consolidation of lexical representations in adults. Finally, I will discuss additional mechanisms via which prediction might play a role in L2 learning, by generating and/or sustaining learners' motivation.

Prediction and L1 vocabulary acquisition in children

The idea that prediction is a key mechanism in L1 acquisition has received considerable attention over the last few years (e.g., Havron et al., 2019; Lindsay et al., 2019; Mani & Huettig, 2012; Rabagliati et al., 2016; Reuter et al., 2019; Saffran, 2020; see also Karaca et al., this volume). Young children are not only sensitive to the degree of expectedness of words in context (e.g., Friedrich & Friederici, 2004), but they are also able to anticipate upcoming referents based on semantic (Borovsky, 2016; Borovsky et al., 2012; Mani et al., 2016; Mani & Huettig, 2012), structural (Gambi et al., 2016; Havron et al., 2019; Lukyanenko & Fisher, 2016) and potentially even phonological information (Gambi, et al., 2018; Mahr et al., 2015; Ylinen et al., 2017). For example, from the age of 3, toddlers are able to combine the meaning of the subject and main verb of a transitive sentence to predict the most likely object (e.g., *pirate* and *chase* predicts *ship*, but *dog* and *chase* predicts *cat*; Borovosky et al., 2012).

But while it is clear that young children are able to predict upcoming linguistic input, the role that these prediction abilities play in language acquisition (if any) is still a matter of debate. One view is that gains in vocabulary knowledge over the course of language acquisition may be a prerequisite for gains in prediction skill (Rabagliati et al., 2016), i.e. prediction skill is a product of linguistic development but does not play any causal role in it. On a different view, prediction skill is one of the driving mechanisms behind linguistic development (Rabagliati et al., 2016). Proponents of the latter view further disagree on the precise nature of this causal role. Below we will discuss two possibilities: (1) that prediction drives learning via the computation of prediction errors and (2) that prediction drives learning by pre-activating familiar words and freeing up resources for the processing of novel words (Fernald et al., 2008; Gambi et al., 2020).

A number of studies reported positive correlations between children's prediction skills and their vocabulary knowledge, even after controlling for age-related increases in vocabulary knowledge (Borovsky et al., 2012; Mani & Huettig, 2012); but see Gambi et al. (2016, 2018) for contrasting findings. These correlations are compatible with the view that prediction skill is a product of development: After all, a child can only use verb semantics (for example) to predict the most likely upcoming argument if they have already acquired the meaning and subcategorization information of the verb (Rabagliati et al., 2016).

But is there any evidence for the proposal that prediction plays a causal role in linguistic development? Theoretically, this proposal derives from ideas in classical learning theory (e.g., Rescorla & Wagner, 1972), where learning is proportional to the discrepancy between expected and observed outcomes. For example, if the learner predicts, on the basis of the combination of cues present in a particular context and of previous experience, that a given outcome is unlikely, observing that outcome will lead to a large discrepancy between the learner's expectations and what they actually observe (i.e., a large prediction error). The result will be a large change to the learner's internal representation of the relation between that context and that outcome. Thus, surprising outcomes lead to larger changes in the learner's model of the environment – i.e., more learning.

This influential idea has been implemented in some models of language acquisition. For example, Ramscar et al. (2013) proposed that children's acquisition of irregular plurals in English can be explained by a model that learns to associate the presence of semantic features (cues) to certain morphological forms (outcomes) via the computation of prediction errors. The model is particularly good at explaining children's overregularization errors (e.g., *mouses* for *mice*): The more children learn frequent, regular plural forms, the more they will predict the occurrence of the regular plural morpheme *-s* in the presence of the semantic feature *plural*; but over time, the accumulation of prediction errors generated by the non-occurrence of the regular plural *mouses* when both the feature *mousiness* and the feature *plural* are present will lead to this form being gradually abandoned. Similarly, several connectionist models of the acquisition of syntactic structure implement the idea that children learn by incrementally attempting to predict the next lexical item in a sequence (Chang, 2002; Chang et al., 2006; Elman, 1990; John & McClelland, 1990), and there is evidence that less expected sentence structures are more likely to be reused (e.g., Peter et al., 2015); I will not discuss syntax further here given the focus of this chapter is on acquisition of vocabulary.

What is the evidence that prediction errors help children learn new words? At first glance, this idea seems rather intuitive. Infants and young children are sensitive to novel, surprising and unexpected information (e.g., Baillargeon et al., 1985; Berger et al., 2006) and the disconfirmation of a prediction (which results in prediction error) could be perceived as surprising by children. Importantly, surprising events do not just command more attention; they also stimulate inquisitive explorative behavior in infants, leading to better learning outcomes (Legare, 2012; Stahl & Feigenson, 2015).

Stahl and Feigenson (2017) set out to test whether witnessing surprising actions would also boost learning of a novel word describing the action in 3-to-6-year-olds.

While children were indeed more likely to remember a novel action word when it referred to an action that was surprising because it violated physical "core knowledge" (e.g., a bag "magically" changing the color of objects that are put inside it), an issue with the control condition makes the findings from this study hard to interpret. In this control condition, no expectation was violated (i.e., the object behaved "normally") and children did not learn the novel word at all (they performed at chance). This suggests that the control condition failed to introduce any salient action that could function as a referent for the novel action word, resulting in a pragmatically infelicitous naming event.

In contrast, Reuter et al. (2019) devised an experiment that allowed for an improved control condition: Rather than comparing the violation of a strong expectation to a no-violation condition, they instead compared two violation conditions that differed only in the strength of the preceding expectation. Three-to-five-year-olds heard a novel pseudoword (e.g., *cheem*), while looking at a visual display with two objects, one familiar (e.g., a spoon) and one unfamiliar (e.g., a butter churner); children typically follow the mutual exclusivity principle (Merriman & Bowman, 1989), and tend to map the pseudoword onto the unfamiliar object. Crucially, Reuter et al. manipulated expectation strength by embedding the pseudoword into sentence contexts that were either more constraining (e.g., *Yummy, let's eat soup. I'll stir it with a … cheem*) or less constraining (e.g., *Neat! Look over there. Take a look at the…cheem*). The more constraining context generated a stronger expectation that the familiar object (*spoon*) would be mentioned, compared to the less constraining context. In both cases, however, this expectation was then violated because the children heard *cheem* instead of a familiar word. Children's eye-movements were recorded while listening to the sentences (to measure expectation strength), and then their ability to map each pseudoword onto the corresponding referent was tested after a short delay (to assess word learning).

Reuter et al. (2019) reasoned that, if prediction error plays a role in children's word learning, a stronger expectation, when violated, should lead to more learning because it generates a larger prediction error. In support of this hypothesis, there was a correlation between children's looking behavior while listening to more constraining contexts and their word learning performance. Specifically, the children who learned more from the more constraining contexts where those who generated stronger expectations (i.e., looked more at the familiar object before hearing the pseudoword), but were then able to revise them once they were violated (i.e., looked less at the familiar object after hearing the pseudoword). This finding suggests that both prediction and revision skills may be important for word learning, and Reuter et al. (2019) argued that their findings support error-driven word learning.

However, Gambi et al. (2021) recently argued that other aspects of Reuter et al.'s findings weaken this conclusion: Children's overall learning performance with pseudowords embedded in more constraining contexts was no better than chance, while they were above chance for pseudowords embedded in *less* constraining contexts. This is at odds with the idea that prediction errors drive children's word learning, because children should experience a larger prediction error following more constraining contexts and these contexts should therefore lead to better, not worse, word learning. In contrast, this finding adds to previous evidence that errors of prediction can in fact be detrimental to children's word learning (Benitez & Saffran, 2018; Benitez & Smith, 2012; Fitneva & Christiansen, 2017). Moreover, in a large-scale conceptual replication of the Reuter et al. study (Gambi et al., 2021), we recently tested 166 2-to-4-year-olds and found that a stronger prediction error did not lead to better word learning. Instead, children remembered novel word-referent mappings at the same rate regardless of the magnitude of prediction errors during learning.

But if larger prediction errors do not facilitate word learning, then what explains the relationship between children's prediction and revision skills and their word learning performance in Reuter et al.'s (2019) study? An alternative interpretation to the one provided by Reuter et al. is as follows: We know that children who are better at revising strong predictions are also more linguistically advanced (Gambi et al., 2020), and that more linguistically advanced children are faster at processing language (Fernald et al., 2006), so perhaps it is these faster language processing skills that account for their superior word learning rather than their ability to revise incorrect predictions. Preliminary support for this hypothesis comes from a recent two-phase study (Gambi et al., 2020). In the first phase of the study, we measured prediction and revision skills and word processing speed in a large sample of 215 2-to-5 year olds using eye-tracking, and correlated each of these skills with children's concurrent receptive vocabulary size. In the second phase of the study, we followed up an opportunistic sub-sample of 55 children to obtain a measure of their vocabulary development over time. While the ability to revise following an incorrect prediction was strongly predictive of children's concurrent vocabulary (over-and-above their age), we found no evidence that it predicted vocabulary development over time – i.e., that it was related to learning.

In contrast, children's ability to predict an upcoming word was related to vocabulary development over time, as was children's speed at processing words. Taken together, these findings suggest that prediction may well play a causal role in vocabulary learning in children, but that this role may be different to the one postulated by error-driven learning accounts. Prediction may not help by generating and revising incorrect expectations, with the resulting prediction error signals driving changes in internal representations; rather, prediction may serve to pre-activate relevant linguistic representations, speeding up processing of familiar words and freeing up resources that children can use to process new, unfamiliar words instead

(Gambi et al., 2020; see also Fernald et al., 2008). In sum, some role for prediction in children's word learning seems likely, but more research is needed to fully understand the underlying mechanisms. In the next section, I will discuss how prediction shapes adult word learning.

Prediction and vocabulary learning in adults

As already mentioned, relatively few studies have tested the role of prediction in adult word learning. Since there is so little evidence, in this section I will draw on studies that tested learning of L2 vocabulary as well as those that tested learning of novel (L1) pseudowords. I will also discuss two studies that did not test learning of new words at all, but rather measured changes in processing of familiar L1 words as a result of prediction during processing of recent linguistic input (Rommers & Federmeier, 2018a, 2018b).

A large body of evidence demonstrates that adult native speakers of a language are skilled predictors (see Pickering & Gambi, 2018, for a recent review). However, alongside clear evidence for what adults can predict, a growing literature shows some limitations to adult prediction: There are important individual differences in predictive ability as a function of (among others) literacy (Mishra et al., 2012) and vocabulary knowledge (Hintz et al., 2017), and increased cognitive load can slow down predictions (Ito et al., 2018).

Of most relevance here, there are differences between prediction in a second language, compared to one's native language (Kaan, 2014; Pickering & Gambi, 2018), even when knowledge of the relevant linguistic features can be demonstrated in L2 comprehenders. This is especially true for predictions at the level of syntax (Foucart et al., 2014; Foucart et al., 2016; Mitsugi & MacWhinney, 2016) and form (Ito et al., 2016; Martin et al., 2013). While it is unclear whether these differences imply that prediction in L2 is limited (e.g., Grüter et al., 2017) or rather that L2 comprehenders' experience of the language is different (Kaan, 2014), these differences may well affect the extent to which prediction drives learning in L2 compared to L1.

With this caveat in mind, is there any evidence that adult word learning (unlike children's) is driven by prediction errors? Interestingly, guessing benefits memory for new L2 words in adults (Potts & Shanks, 2014; Grüter et al., this volume). Specifically, when participants are tasked with guessing what an unfamiliar L2 word means before being told its correct meaning, they are more likely to remember that meaning compared to simply studying it or choosing a possible meaning out of a list without generating a guess first. One possible reason why guessing helps is that it is likely to result in errors, and the comparison between the incorrect guess and the correct meaning will generate a strong prediction error. However, there are other likely reasons too. First, the very act of generating a guess may trigger a sense of curiosity to know the correct answer, and thus lead to more active processing of that answer (see the section *Prediction and motivation in L2 vocabulary learning*). In addition, participants rate the guessing condition as more difficult than other learning conditions, suggesting that it may be an example of those "desirable difficulties" that foster deeper, longer-lasting learning (Bjork & Kroll, 2015).

More direct evidence in support of error-driven word learning comes from the finding that adults perform better in a cross-situational word learning task (Yu & Smith, 2007) if they are exposed to a greater proportion of unexpected word-referent mappings (Fitneva & Christiansen, 2011, 2017). In these studies, unexpectedness was manipulated artificially by changing a subset of the mappings between novel words and referents (unbeknownst to the participants), following an initial familiarization phase. For example, the pseudoword *dipe* may have initially been paired with the picture of one kind of orchid, but it would later be paired with a different kind of orchid. Strikingly, when the proportion of unexpected (i.e., changed) to expected (i.e., unchanged) mappings was higher, adults actually learned more compared to when it was lower. Fitneva and Christiansen (2011, 2017) observed a learning advantage for both unexpected and expected mappings, which may indicate that participants allocated more attentional resources to the task when the likelihood of an error was higher, rather than benefitting directly from having their expectations disconfirmed. If the latter were true, then the learning benefit should be specific to unexpected mappings and should not extend to expected mappings.

Interestingly, while Fitneva and Christiansen found comparable benefits for expected and unexpected mappings, a subsequent study (Grimmick et al., 2019) using a very similar design instead found that memory was enhanced specifically for unexpected mappings. Grimmick et al. showed that a specific advantage for unexpected mappings cannot be explained by a computational model that includes attentional biases without the addition of an error-driven learning mechanism. But despite this recent finding, on the whole it is still unclear to what extent prediction errors play a role in adult word learning: Most studies are compatible with alternative explanations, and there is still little direct evidence for error-driven word learning.

Thus, we recently set out to provide a direct test of this idea. Gambi et al. (2021) tested both adults and children on a paradigm similar to the one used by Reuter et al. (2019) with children. Recall that this paradigm allows us to compare the effects on word learning of disconfirmed predictions that differ in strength. Furthermore, the strength of these predictions is manipulated by using context sentences which, in conjunction with the visual context, make mention of one particular referent more or less predictable. This way it is possible to implement a manipulation of predictability that is close to that used by countless psycholinguistic experiments on online language processing (Pickering & Gambi, 2018), and examine its "downstream" consequences on learning.

Crucially, adults in our study benefitted from having a stronger (compared to a weaker) prediction disconfirmed: They were more likely to remember a newly-formed association between a novel pseudoword and an unfamiliar object when the pseudoword was presented after a more constraining context (e.g., *Peppa will eat the…cheem.*, when an apple is present on the screen) compared to a less constraining context (e.g. *Peppa will get the…cheem*). This suggests that stronger predictions, when disconfirmed, lead to the formation of stronger memory traces in adults (see also Greve et al., 2017).

Importantly, we also ruled out an alternative explanation: More constraining verbs (e.g., *eat, drive*) are semantically richer than less constraining verbs (e.g., *get, move*), and may benefit memory because they provide richer contextual semantic cues that become associated with the pseudoword. We know that adults can infer the meaning of novel pseudowords from context even after a single exposure (e.g., Borovsky et al., 2010), and most likely they do so by quickly establishing associations between these words and the sentential context they appear in, including perhaps the semantic features pre-activated by the verb (Federmeier & Kutas, 1999). To show this could not account for the memory benefit following more constraining contexts, in a control experiment all sentence contexts used semantically richer verbs (e.g., *eat*), but these biased expectations towards the familiar referent only when this referent fit their constraint (e.g., apple for *eat*) not when a different familiar object was used instead (e.g., car), resulting in larger prediction errors in the former than the latter condition. Despite identical semantic cues across conditions, we again found that larger prediction errors led to better memory for the novel pseudoword, suggesting a role for prediction error in adult word learning (Gambi et al., 2021).

These findings are important for two reasons. First, they paint a striking contrast between the role of prediction error in child vs. adult word learning. At present, it is unclear what is driving this apparent developmental discontinuity. It could be that children compute prediction errors, but not efficiently enough for these to affect their memory for word-referent mappings. Or it could be that children's learning is less sensitive to linguistic prediction errors because of their limited linguistic knowledge: Since their predictions are often disconfirmed by the input, it may not be adaptive for them to concentrate resources on encoding unexpected linguistic input (as most input is unexpected to some degree to them).

Second, these findings are important because they represent a first step towards linking the vast literature on prediction during sentence processing with the literature on adult word learning. An important next step would be to flesh out how and to what extent prediction could support adult learning during naturalistic exposure to a second language (e.g., while watching films, reading newspapers, or while having conversations), as opposed to during vocabulary learning tasks. One key question that future work should address is also how prediction error would support consolidation of partial word knowledge across repeated encounters with the same novel word.

But while a shift towards more naturalistic paradigms should foster new insights into the relationship between processing and learning, at present the little evidence available on this relationship comes from adult native speakers processing their L1. Two recent studies investigated how varying the predictability of words in sentence contexts affected downstream memory access for those words (Rommers & Federmeier, 2018a, 2018b). Rommers and Federmeier (2018b) showed that more predictable words are processed more shallowly than less predictable words: They measured the size of the repetition priming effect – that is, the reduction in the amplitude of the N400 when a word is encountered a second time – and found less repetition priming when the first encounter with a word had been within a context that made that word more predictable. Encountering the word in a more predictable context also led to a reduction of the repetition effect on the magnitude of the Late Positive Component (LPC), which may indicate that participants were less likely to explicitly recall having read the word via the episodic memory system.

In sum, words that were less expected were encoded more strongly in memory, possibly via both implicit and explicit memory processes. However, when participants' memory was tested at the end of the experiment, participants were equally good at recognizing more and less predictable words (Rommers & Federmeier, 2018b). It may be that recognition memory is not affected by predictability, or that predictability both facilitates and hinders recognition memory because predictable words are first pre-activated but then processed more shallowly. One important open question is whether predictability would facilitate or hinder other aspects of memory, such as the ability to recall the context given the target word, which is likely implicated in acquiring the meaning of unfamiliar words from context (Borovsky et al., 2010).

It is also important to ask what happens to the representations of words that are predicted but not actually encountered (because a strong prediction is disconfirmed). Based on findings reviewed above, encountered but unexpected words should be strongly encoded in memory, but how does pre-activation of a word that is never encountered affect downstream processing? Rommers and Federmeier (2018a) used the reduction in N400 amplitude on second presentation of a word to show that words that had been merely expected but not seen still elicited a repetition effect, although this was not as strong as for words that were actually seen twice. However, words that had merely been expected but not seen did not elicit a repetition effect on the magnitude of the LPC and were not recognized any better than words seen just once in an explicit memory task. This suggests that expected words that are not encountered are unlikely to interfere with explicit memory processes, but may well affect implicit memory processes (e.g., priming). This may help elucidate the mechanism by which the meaning of novel words is inferred from predictive contexts.

To conclude, there is still only tentative evidence for a role of prediction, and in particular disconfirmed predictions, in long-term memory consolidation of novel words and their meanings in adults. More studies looking at the processing of novel words in predictable and less predictable contexts are needed and, crucially, these studies should measure the downstream effects of such processing on memory, to uncover the cognitive mechanisms involved. This section focused on the predictability of words in sentence contexts. The next section takes a broader view, to examine how prediction relates to motivation in the context of L2 vocabulary learning.

Prediction and motivation in L2 vocabulary learning

Learning scientists have long recognized that learners benefit from generating expectations – for example guessing the answer to a question, or the solution to a problem – compared to passively absorbing information (e.g., Brod et al., 2018; Pressley et al., 1992; Slamecka & Graf, 1978). In the previous section, we discussed a study by Potts and Shanks (2014) showing this "guessing benefit" applies to memory for L2 vocabulary as well. But as we also highlighted in that section, the precise mechanisms underlying this effect are still unclear and likely to be multifaceted. Crucially, the notion that prediction errors benefit learning may be insufficient to explain this finding for two reasons.

First, the idea that learning is driven by prediction errors fails to explain why guessing is only beneficial to memory if it precedes presentation of the correct answer (Potts et al., 2019): When participants were asked what they would have guessed the meaning of an L2 word to be *after* being told the correct meaning, they were no better at remembering the correct answer than when they were simply asked to study the correct meaning. Note that participants' post-hoc guesses were still very unlikely to be correct and indeed had only a very weak semantic relation to the correct meaning, suggesting the participants were not unduly influenced by knowing the correct answer when generating their guesses. Therefore, this condition should still have generated larger predictions errors compared to making no guess at all; and yet, it did not result in any memory benefit.

So, what does account for the "guessing benefit"? Potts et al. (2019) suggest that guessing leads to deeper encoding of the correct answer because it triggers curiosity, i.e., an intrinsic motivation to know the correct meaning of the L2 word. Accordingly, they showed that self-reported curiosity ratings were higher when they were made after than before generating a guess. Intrinsic motivation, or curiosity, has been researched for decades as a driver of learning (for recent reviews, see Kidd & Hayden, 2015; Murayama et al., 2019), leading to many different theoretical frameworks, but here I will focus on one recent framework (Murayama, 2019) grounded in neuroscientific findings from the last two decades (Gruber & Ranganath, 2019). According to Murayama (2019), the key feature of curiosity-driven learning is that it is motivated by the intrinsic reward value of knowledge. This brings us to the second reason why prediction errors seem insufficient to explain the "guessing benefit" in learning in general and in word learning in particular: Prediction errors do not take into account the reward value of information.

The role of extrinsic rewards in driving learning is well understood: Animals and humans learn to choose actions that will lead to larger rewards (e.g., food, money); the expected reward value of choices is represented (and updated) by a dedicated brain network including dopaminergic areas of the mid-brain and the striatum (e.g., O'Doherty, 2004). Importantly, these same areas show increased activation when participants are placed in a state of heightened curiosity by having them read trivia questions (Kang et al., 2009); moreover, when participants are placed in such a heightened curiosity state, they are more likely to remember not just the answer to the trivia question, but also incidental information presented to them after the trivia question but before the answer is revealed to them (Gruber et al., 2014). Such enhanced memory performance is related to activation of the reward network and its influence on the hippocampus via the release of dopamine (Gruber & Ranganath, 2019). In sum, neuroscientific studies support the idea that a desire for knowledge and information can enhance memory and learning via internally-generated reward signals.

Is there any evidence that reward signals play a role in word learning? Ripollés et al. (2014) showed that successful inference of the meaning of a foreign word from context is associated with increased activation in the ventral striatum, which is part of the reward network, and also with enhanced functional and structural connectivity between the ventral striatum and language-processing areas of the brain. Further, Ripollés et al. (2016) found enhanced functional connectivity between the ventral striatum, hippocampus and dopaminergic areas of the mid-brain in conjunction with successful word learning.

But while these findings suggest that successful word learning can be an intrinsically rewarding experience, it is less clear whether they implicate a causal role for intrinsic rewards in facilitating language learning. Evidence for the latter was instead provided by Ripollés et al. (2018), who showed that participants completing the word learning task under the effect of a dopamine receptor antagonist learned and remembered fewer words. Thus, dopamine does appear to be causally implicated in learning word meanings from context in this task, suggesting that the experience of reward may be crucial for long-term consolidation of newly acquired linguistic information.

What does this mean for theories of L2 vocabulary learning? I propose that these findings highlight the need to incorporate the reward value of linguistic knowledge into cognitive theories of L2 acquisition. The role of motivation in learning a second language has long been researched in the field of second language acquisition (SLA) (e.g., Dörnyei, 2014; Gardner, 1985), but it has largely been overlooked in psycholinguistics. This is a significant gap, because motivation predicts effort and attainment in L2 learning, in both lab-based and classroom studies (Masgoret & Gardner, 2003). Conversely, lack of motivation is the single most frequently cited barrier to learning an additional language in adulthood; for example, in 2012 34% of Europeans said they were discouraged from learning another language because they were not motivated enough to do so, ahead of lack of time or difficulty accessing language courses (European Commission, 2012).

However, it is difficult to link traditional SLA accounts of how motivation affects L2 acquisition to psycholinguistic theories about language learning. While this is not the place to review the SLA accounts in detail, many incorporate the concept of self-efficacy – that is, the sense of being in control of one's own learning. The key proposal is that self-efficacy is greater when learners are intrinsically motivated (rather than motivated by external pressures or rewards), and that learners who feel in control are more likely to invest time and resources in the learning process, compared to learners who feel controlled by external forces (Deci et al., 1999). Interestingly, motivated learners are thought not just to spend more time on learning tasks or exposing themselves to L2 input: There is some evidence they are also more likely to adopt a larger number of learning strategies and be more aware of how they learn (e.g., Bonney et al., 2008).

But crucially, it is an open question whether motivation alters the basic cognitive mechanisms of second language learning. The reward framework for intrinsic motivation (Murayama, 2019) may help bridge this gap because it suggests a set of mechanisms (some of which have known neural correlates) that account for how motivation affects learning and how learning in turn affects the development and maintenance of motivation. In the remainder of this section I will outline three ways in which prediction mechanisms could be implicated in this two-way relationship between intrinsic rewards and learning.

First, reward-learning itself depends upon the computation of predictions and prediction errors, though crucially these predictions do not relate to future states of the world but rather to the occurrence and magnitude of future rewards (O'Doherty, 2004). What does this mean in the context of language learning? One possibility is that knowledge about a second language is in itself a reward: The work by Ripollés et al. (2014, 2016, 2018) appears to support this idea and, anecdotally, many language learners derive pleasure from mastery of the language itself. If this is the case, then predictions about upcoming linguistic input (e.g., the meaning of an unfamiliar word) and predictions about future rewards might be tightly linked.

However, language understanding is not always an aim in itself for learners. Many learners are motivated by what they could do with knowledge of the language: Consuming the culture associated with the language, communicating with other speakers of the language, or finding a job (Dörnyei, 2014). In all of these cases, reward predictions would still be related to linguistic predictions (because linguistic predictions support understanding), but the relationship would be much more complex and indirect. Thus, it is an open question how predictions about these communicative rewards would affect the process of learning from linguistic prediction errors.

Second, one important feature of curiosity-driven learning is its self-sustaining nature (Murayama, 2019): New information is pursued because it has the potential to close a gap in one's knowledge, and the accumulation of information in turn leads to the discovery of new knowledge gaps that were not apparent before, thus driving further learning. Importantly, according to this information-gap theory of curiosity (Loewenstein, 1994), prediction plays a key role in highlighting the presence of an information gap. Surprise is often the starting point for detection of an information gap, so generating expectations based on current knowledge, and finding them to be disconfirmed by the evidence can help identify a knowledge gap. Alternatively, if an expectation can be generated, but it is associated with a certain degree of uncertainty (ideally, neither too little nor too much uncertainty; Kidd & Hayden, 2015), this can stimulate a search for further evidence to help confirm or disconfirm that expectation. In sum, disconfirmed or uncertain predictions can help kickstart reward-driven learning.

Third, there is emerging evidence that one of the ways in which being in state of heightened curiosity benefits memory is by sharpening learning from prediction errors. Murty and Adcock (2014) had participants complete a color change-detection task under conditions of low motivation (with low monetary rewards) or high motivation (with high monetary rewards); the reward value varied trial-by-trial and was indicated by a cue at the beginning of the trial. Crucially, after the cue, but before the target, they sometimes presented an unexpected stimulus; the stimulus was unexpected because it differed subtly from a series of repeating stimuli that preceded it, but participants were required to ignore it. At the end of the session, participants' recognition memory for the unexpected stimuli was tested and found to be greater than chance only in the high motivation condition. Moreover, fMRI data showed that sensitivity to unexpected stimuli in a small region of the hippocampus was enhanced under conditions of high motivation, and that this enhancement was related to increased activity in the ventral tegmental area, which is part of the dopaminergic mid-brain involved in reward-driven learning. Thus, these findings suggest that being in state of heightened motivation increases sensitivity to learning from unexpected events, and memory for these events (Shohamy & Adcock, 2010).

A key open question is thus whether motivation to learn a language would similarly change sensitivity to learning from linguistic prediction errors. If so, this would provide a mechanistic explanation for the positive effect of motivation on language attainment above and beyond increased effort and exposure to input. We already know that comprehenders can flexibly adapt their predictions (Yurovsky et al., 2017) and the extent to which they rely on pre-activated information vs. bottom-up evidence (Gibson et al., 2013) depending on task goals and the statistics of the linguistic input (see also Kuperberg & Jaeger, 2016), but we are not aware of any study that has investigated whether motivation can affect the extent to which L2 learners are sensitive to linguistic prediction errors.

However, some suggestive evidence comes from a recent study, where 2-year-olds learned novel word-referent associations (Ackermann et al., 2020). Prior to learning, each infant's interest in (i.e., curiosity for) the referents and the categories they belonged to (e.g., animals vs. drinks) was measured using changes in pupil dilation as an index of interest (Mathôt, 2018). Infants' success at word learning was related to their interest in both specific referents and their categories. However, these findings are correlational, so it is unclear whether interest plays a causal role in learning. It is also unclear how interest affects learning (i.e., via what mechanism). And crucially, the study did not manipulate the expectedness of words, so we do not know whether increased interest benefits word learning by augmenting sensitivity to prediction errors.

Summary and open questions

What is the role of prediction in L2 vocabulary learning? In this chapter, I began by reviewing what we know about the role of prediction in L1 vocabulary development during childhood to highlight two different causal pathways: Prediction may facilitate learning by pre-activating representations for familiar words, thereby freeing up cognitive resources for processing novel words and/or prediction might drive learning via the computation of prediction errors. While this is largely still an open question in the field of first language acquisition, evidence is emerging that the role of prediction errors in vocabulary development might be somewhat limited. In contrast, we have some evidence that prediction errors play a role in encoding and consolidation of novel words in memory for adult learners. But we still know very little about the "downstream" consequences of confirmed and disconfirmed predictions on long-term linguistic representations. Finally, I have argued that predictions about upcoming linguistic input are unlikely to be the only prediction-based mechanism underlying L2 vocabulary learning: Cognitive theories of L2 vocabulary learning should also incorporate predictions about the reward value of linguistic knowledge, so that we can begin to understand how prediction mechanisms can help generate and sustain motivation to learn a second language. Key questions for this line of research will be: Is linguistic knowledge rewarding in itself or because it allows successful communication? How does prediction foster curious and active learning of L2 vocabulary? Does motivation modulate learners' sensitivity to linguistic prediction errors?

Acknowledgments

Many of the ideas discussed here were developed in joint work with Hugh Rabagliati and Martin Pickering, but any mistakes or inaccuracies remain my own. I would also like to thank Matthias Grüber for introducing me to the science of curiosity and for many useful discussions.

Funding

Writing of this chapter was partly supported by a British Academy Small Grant to Chiara Gambi and Kate Messenger (SRG1920_100600).

References

Ackermann, L., Hepach, R., & Mani, N. (2020). Children learn words easier when they are interested in the category to which the word belongs. *Developmental Science*, 23(3), e12915. https://doi.org/10.1111/desc.12915

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20(3), 191–208. https://doi.org/10.1016/0010-0277(85)90008-3

Benitez, V. L., & Saffran, J. R. (2018). Predictable events enhance word learning in toddlers. *Current Biology*, 28(17), 2787–2793. https://doi.org/10.1016/j.cub.2018.06.017

Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition*, 125(3), 339–352. https://doi.org/10.1016/j.cognition.2012.08.006

Berger, A., Tzur, G., & Posner, M. I. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences*, 103(33), 12649–12653. https://doi.org/10.1073/pnas.0605350103

Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *The American Journal of Psychology*, 128(2), 241–252. https://doi.org/10.5406/amerjpsyc.128.2.0241

Bonney, C. R., Cortina, K. S., Smith-Darden, J. P., & Fiori, K. L. (2008). Understanding strategies in foreign language learning: Are integrative and intrinsic motives distinct predictors? *Learning and Individual Differences*, 18(1), 1–10. https://doi.org/10.1016/j.lindif.2007.11.005

Borovsky, A. (2016). The amount and structure of prior event experience affects anticipatory sentence interpretation. *Language, Cognition and Neuroscience*, 32(2), 190–204. https://doi.org/10.1080/23273798.2016.1238494

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, 112(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, 116(2), 289–296. https://doi.org/10.1016/j.cognition.2010.05.004

Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: The element of surprise. *Learning and Instruction*, 55, 22–31. https://doi.org/10.1016/j.learninstruc.2018.01.013

Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, 26(5), 609–651. https://doi.org/10.1207/s15516709cog2605_3

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668. https://doi.org/10.1037/0033-2909.125.6.627

Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge. https://doi.org/10.4324/9781410613349

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211. https://doi.org/10.1207/s15516709cog1402_1

European Commission (2012). *Special Eurobarometer 386: Europeans and their languages*. Retrieved from https://www.eubusiness.com/Books/languages-12

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41(4), 469–495. https://doi.org/10.1006/jmla.1999.2660

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98

Fernald, A., Marchman, V. A., & Hurtado, N. (2008). Input affects uptake: How early language experience influences processing efficiency and vocabulary learning. In *Proceedings of the 7th IEEE International Conference on Development and Learning, Monterey, CA* (pp. 37–42). https://doi.org/10.1109/DEVLRN.2008.4640802

Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate word learning. *Cognitive Science*, 35(2), 367–380. https://doi.org/10.1111/j.1551-6709.2010.01156.x

Fitneva, S. A., & Christiansen, M. H. (2017). Developmental changes in cross-situational word learning: The inverse effect of initial accuracy. *Cognitive Science*, 41(S1), 141–161. https://doi.org/10.1111/cogs.12322

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1461–1469. https://doi.org/10.1037/a0036756

Foucart, A., Ruiz-Tada, E., & Costa, A. (2016). Anticipation processes in L2 speech comprehension: Evidence from ERPs and lexical recognition task. *Bilingualism: Language and Cognition*, 19(01), 213–219. https://doi.org/10.1017/S1366728915000486

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, 16(8), 1465–1477. https://doi.org/10.1162/0898929042304705

Gambi, C., Gorrie, F., Pickering, M. J., & Rabagliati, H. (2018). The development of linguistic prediction: Predictions of sound and meaning in 2- to 5-year-olds. *Journal of Experimental Child Psychology*, 173, 351–370. https://doi.org/10.1016/j.jecp.2018.04.012

Gambi, C., Jindal, P., Sharpe, S., Pickering, M. J., & Rabagliati, H. (2020). The relation between preschoolers' vocabulary development and their ability to predict and recognize words. *Child Development*. https://doi.org/10.1111/cdev.13465

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond Associations: Sensitivity to structure in pre-schoolers' linguistic predictions. *Cognition*, 157, 340–351. https://doi.org/10.1016/j.cognition.2016.10.003

Gambi, C., Pickering, M. J., & Rabagliati, H. (2021). Prediction error boosts retention of novel words in adults but not in children. *Cognition*, 211. https://doi.org/10.1016/j.cognition.2021.104650

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Edward Arnold.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, 110(20), 8051–8056. https://doi.org/10.1073/pnas.1216438110

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, 94, 149–165. https://doi.org/10.1016/j.jml.2016.11.001

Grimmick, C., Gureckis, T. M., & Kachergis, G. (2019). Evidence of error-driven cross-situational word learning. Paper presented at the 41st Annual Conference of the Cognitive Science Society, Montreal, Canada. https://cogsci.mindmodeling.org/2019/papers/0084

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, 84(2), 486–496. https://doi.org/10.1016/j.neuron.2014.08.060

Gruber, M. J., & Ranganath, C. (2019). How curiosity enhances hippocampus-dependent memory: The Prediction, Appraisal, Curiosity, and Exploration (PACE) framework. *Trends in Cognitive Sciences*, 23(12), 1014–1025. https://doi.org/10.1016/j.tics.2019.10.003

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, 7(2), 199–229. https://doi.org/10.1075/lab.15011.gru

Havron, N., de Carvalho, A., Fiévet, A. C., & Christophe, A. (2019). Three-to four-year-old children rapidly adapt their predictions and use them to learn novel word meanings. *Child Development*, 90(1), 82–90. https://doi.org/10.1111/cdev.13113

Hintz, F., Meyer, A. S., & Huettig, F. (2017). Predictors of verb-mediated anticipatory eye movements in the visual world. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(9), 1352–1374. https://doi.org/10.1037/xlm0000388

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism: Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Ito, A., Martin, A. E., & Nieuwland, M. S. (2016). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience*, 32(8), 954–965. https://doi.org/10.1080/23273798.2016.1242761

John, M. F. S., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, 46(1–2), 217–257. https://doi.org/10.1016/0004-3702(90)90008-N

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-Y., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, 20(8), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, 88(3), 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

Kuperberg, G. R., & Jaeger, T. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Legare, C. H. (2012). Exploring explanation: Explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development*, 83(1), 173–185. https://doi.org/10.1111/j.1467-8624.2011.01691.x

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychological Science*, 30(4), 504–515. https://doi.org/10.1177/0956797618822802

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition*, 146, 349–370. https://doi.org/10.1016/j.cognition.2015.10.012

Mahr, T., McMillan, B. T., Saffran, J. R., Weismer, S. E., & Edwards, J. (2015). Anticipatory coarticulation facilitates word recognition in toddlers. *Cognition*, 142, 345–350. https://doi.org/10.1016/j.cognition.2015.05.009

Mani, N., Daum, M. M., & Huettig, F. (2016). "Pro-active" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. *Quarterly Journal of Experimental Psychology*, 69(11), 2189–2201. https://doi.org/10.1080/17470218.2015.1111395

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843–847. https://doi.org/10.1037/a0029284

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53(S1), 167–210. https://doi.org/10.1111/1467-9922.00227

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, 1(1), 16. https://doi.org/10.5334/joc.18
Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, 54, 1–132. https://doi.org/10.2307/1166130
Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1), 1–10. https://doi.org/10.16910/jemr.5.1.3
Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(01), 19–35. https://doi.org/10.1017/S1366728914000881
Murayama, K. (2019). *A reward-learning framework of autonomous knowledge acquisition: An integrated account of curiosity, interest, and intrinsic-extrinsic rewards* (Unpublished manuscript). Retrieved from: https://osf.io/zey4k
Murayama, K., FitzGibbon, L., & Sakaki, M. (2019). Process account of curiosity and interest: A reward-learning perspective. *Educational Psychology Review*, 31, 875–895. https://doi.org/10.1007/s10648-019-09499-9
Murty, V. P., & Adcock, R. A. (2014). Enriched encoding: Reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cerebral Cortex*, 24(8), 2160–2168. https://doi.org/10.1093/cercor/bht063
O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, 14(6), 769–776. https://doi.org/10.1016/j.conb.2004.10.016
Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, 81, 1–15. https://doi.org/10.1016/j.jml.2014.12.002
Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158
Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(4), 329–392. https://doi.org/10.1017/S0140525X12001495
Potts, R., Davies, G., & Shanks, D. R. (2019). The benefit of generating errors during learning: What is the locus of the effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(6), 1023–1041. https://doi.org/10.1037/xlm0000637
Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, 143(2), 644–667. https://doi.org/10.1037/a0033194
Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27(1), 91–109. https://doi.org/10.1207/s15326985ep2701_7
Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition, and Neuroscience*, 31(1), 94–105. https://doi.org/10.1080/23273798.2015.1077979
Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language*, 89(4), 760–793. https://doi.org/10.1353/lan.2013.0068

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64–99.

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*, 55(8), 1656–1665. https://doi.org/10.1037/dev0000754

Ripollés, P., Ferreri, L., Mas-Herrero, E., Alicart, H., Gómez-Andrés, A., Marco-Pallares, J., Antonijoan, R. M., Noesselt, T., Valle, M., Riba, J., & Rodríguez-Fornells, A. (2018). Intrinsically regulated learning is modulated by synaptic dopamine signaling. *Elife*, 7, e38113. https://doi.org/10.7554/eLife.38113

Ripolles, P., Marco-Pallares, J., Alicart, H., Tempelmann, C., Rodríguez-Fornells, A., & Noesselt, T. (2016). Intrinsic monitoring of learning success facilitates memory encoding via the activation of the SN/VTA-Hippocampal loop. *Elife*, 5, e17441. https://doi.org/10.7554/eLife.17441

Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H.-J., Rodríguez-Fornells, A., & Noesselt, T. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology*, 24(21), 2606–2611. https://doi.org/10.1016/j.cub.2014.09.044

Rommers, J., & Federmeier, K. D. (2018a). Lingering expectations: A pseudo-repetition effect for words previously expected but not presented. *NeuroImage*, 183, 263–272. https://doi.org/10.1016/j.neuroimage.2018.08.023

Rommers, J., & Federmeier, K. D. (2018b). Predictability's aftermath: Downstream consequences of word predictability as revealed by repetition effects. *Cortex*, 101, 16–30. https://doi.org/10.1016/j.cortex.2017.12.018

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, 14(1), 49–54. https://doi.org/10.1111/cdep.12355

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, 14(10), 464–472. https://doi.org/10.1016/j.tics.2010.08.002

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, 348(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Stahl, A. E., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*, 163, 1–14. https://doi.org/10.1016/j.cognition.2017.02.008

Ylinen, S., Bosseler, A., Junttila, K., & Huotilainen, M. (2017). Predictive coding accelerates word recognition and learning in the early stages of language development. *Developmental Science*, 20(6), e12472. https://doi.org/10.1111/desc.12472

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yurovsky, D., Case, S., & Frank, M. C. (2017). Preschoolers flexibly adapt to linguistic input in a noisy channel. *Psychological Science*, 28(1), 132–140. https://doi.org/10.1177/0956797616668557

CHAPTER 10

Forcing prediction increases priming and adaptation in second language production

Theres Grüter[1], Yanxin Alice Zhu[1] and Carrie N. Jackson[2]
[1]University of Hawaiʻi at Mānoa / [2]Penn State University

This paper presents a priming experiment with Korean learners of English designed to test the hypothesis that engagement in prediction and the consequent computation of prediction error will lead to increased priming and adaptation. Participants in the guessing-game condition, who had to predict a virtual partner's description of prime pictures, showed marginally greater immediate priming and significantly greater adaptation in terms of change from baseline to post-test than those in a standard repetition priming condition, consistent with error-based learning accounts of structural priming. Effects were largest for learners in the middle of the proficiency range. Findings from this study suggest that priming best facilitates L2 learning when learners are engaged in the proactive creation of expectations about upcoming information.

Introduction

Our success in daily life is critically dependent on our ability to adapt to our environment and to learn from our mistakes. This is evident in our successful use of language in everyday communication, as well as in our ability to learn a language from processing the linguistic and non-linguistic information in the input we experience. Traditionally, these two abilities – language production and comprehension on the one hand, and language learning and development on the other – were the focus of largely separate academic fields, with psycholinguists attending to language production and comprehension mechanisms in (mostly) native adult language users, and applied linguists and developmental psychologists attending to language learning. More recently, approaches from the broader field of cognitive science have sought to account for properties of language processing and learning via a set of more unified cognitive mechanisms. In particular, the view of language processing as driven, at least in substantial part, by prediction (Clark, 2013; Pickering & Gambi, 2018), and prediction error constituting a driving force in learning (Chang et al., 2006; Jaeger

https://doi.org/10.1075/bpa.12.10gru

& Snider, 2013), has presented an exciting framework for unifying the study of language processing and language learning. The extension of this perspective to the field of Second Language Acquisition (SLA) is still in its infancy (but see Jackson & Hopp, 2020; Kaan & Chun, 2018; for further discussion, Hopp, this volume), yet we believe holds great promise not only for increasing our understanding of the mechanisms that underlie language learning after early childhood, but also for a better integration of SLA in the wider field of cognitive science.

The study we present in this paper was motivated by the proposal that prediction error drives (second language) learning. Within this context, we define *prediction* broadly as expectations about information that has not yet been encountered. Within broader theoretical frameworks of predictive coding in cognition (e.g., Bar, 2007; Clark, 2013), prediction is assumed to operate largely outside of conscious awareness. However, the relation between prediction and conscious awareness, and in particular the extent to which attention and awareness guide the formation of expectations remains largely unknown (Meijs et al., 2018), and was listed among key questions for future research by Bar (2007). Against this backdrop, we deliberately adopt a broad definition of prediction that includes presumably implicit, contextually conditioned pre-activation of upcoming words in a sentence, as typically examined in psycholinguistic experiments on predictive sentence processing (e.g., Pickering & Gambi, 2018), as well as more explicit guessing about unknown information, as in the learning of new words (e.g., Potts et al., 2019). This broad definition is further motivated by our focus being not on the nature of prediction itself but on the downstream consequences of comparing predicted with actually encountered information, i.e., the computation of prediction error. Error-based learning has been a topic of relatively recent interest in the psycholinguistic literature on sentence processing, but has a long-standing history in learning theory (e.g., Rescorla & Wagner, 1972; see also Gambi, this volume). Drawing on insights from across these literatures, we hypothesized that if second language (L2) learners are put in a situation where they are forced to predict upcoming linguistic input, and forced to attend to potential discrepancies between their prediction and the actual input they then encounter, they will be more likely to use the encountered structures in their own subsequent production. To test this hypothesis, we conducted a written production priming experiment, focusing on double-object datives, with Korean L2 learners of English. The results show that participants who were forced to predict exhibited marginally greater immediate priming and significantly greater adaptation in terms of change from baseline to post-test, than participants who followed a standard repetition priming procedure.

Structural priming and learning from prediction error

Research on structural priming in language production and comprehension has shown that language users are more likely to use a particular syntactic structure when they have recently encountered that structure (Bock, 1986; Pickering & Ferreira, 2008). Such alignment with recent input was originally attributed to increased activation of the structure in question, making it temporarily more accessible to the production and comprehension systems (Pickering & Branigan, 1998). Residual activation accounts provide a straightforward explanation for immediate priming, that is, increased use of a primed structure in an immediately following target trial. Yet these accounts are less able to capture the observation that priming can persist over longer and variable time intervals (e.g., Bock & Griffin, 2000; Kaschak et al., 2014). Such long(er) term priming effects that persist beyond immediately adjacent trials gave rise to alternative accounts that attribute these effects to mechanisms of implicit learning (Chang et al., 2006; Dell & Chang, 2014). In this paper, we will refer to such longer term priming effects as *adaptation*; we use the term *immediate priming* for effects in immediately adjacent trials.

While the specifics of the proposed learning mechanisms remain a matter of on-going debate (Myslín & Levy, 2016), an assumption common to all implicit learning accounts is that it is error-driven (Dell & Chang, 2014; Jaeger & Snider, 2013). In other words, the adaptive behavior that we observe is assumed to be the result of the comprehender perceiving a mismatch between a structure they encountered in the input and the structure they would have expected in that context given their knowledge and previous experience. For example, if upon observing lightning hit a church, a speaker says *That church was just hit by lightning*, the use of a passive will cause a comprehender to experience a certain amount of surprisal given that it is generally more common to describe transitive events in the active voice. Due to this mismatch between an observed structure and one that would have been more likely given the comprehender's previous experience, the comprehender's bias for expecting and producing passives in the future will be adjusted in the direction of increased likelihood of passives. The benefit of this adjustment lies in the decrease of surprisal and disruption the comprehender will experience the next time they encounter a passive to describe a transitive event. The size of the adjustment will be a function of the amount of surprisal experienced: the more unexpected the encountered structure, the greater the adjustment. This is known as the "inverse frequency effect," which is reflected in greater priming for infrequent versus frequent constructions (Pickering & Ferreira, 2008).

The overall goal of error-driven learning is to develop a system in which expectations optimally match the input and minimize surprisal. The hypothesized adjustments to the system due to prediction error can be captured in terms of adjusted weights in a recurrent network, or a shift in distributional expectations conditioned by surprisal within a Bayesian belief update model (for further discussion see Myslín & Levy, 2016). A critical assumption underlying all models of learning by prediction error is that language users generate predictions. This assumption seems uncontroversial in the context of broader claims that human brains are in essence "prediction machines," and that prediction is a pervasive property of human cognition and behavior far beyond language processing (Clark, 2013). Recent work has shown, however, that the contribution of top-down anticipatory processes relative to bottom-up integrative processes in various cognitive domains, including language comprehension, can vary substantially depending on multiple factors that are still not well understood. In the context of spoken language processing by adult native speakers, for example, participant-level factors such as age (e.g., Wlotko et al., 2012) and literacy (e.g., Mishra et al., 2012; for review, see Huettig & Pickering, 2019) have been shown to modulate language users' reliance on prediction. Similarly, task-related factors, such as the amount of preview time in visual-world experiments (Ferreira et al., 2013; Sorenson & Bailey, 2007) or the rate at which linguistic stimuli are presented (Ito et al., 2017), can modulate engagement in predictive processing. Such observations have led to on-going debate on the extent to which prediction is a driving, or even a necessary, force in language processing (Huettig, 2015; Ferreira & Chantavarin, 2018).

This view of prediction as a potentially less pervasive and more variable force in language processing has important consequences for accounts of language learning based on the computation of prediction error. If participant- and task-related factors modulate reliance on prediction, then variability in these factors should be predictive (no pun intended) of the amount of learning that occurs. In other words, reduced engagement in prediction, and thus reduced experience of prediction error, should lead to reduced learning. For implicit learning accounts of priming, this leads to the hypothesis that the size of priming effects will be modulated by participant- and task-related variability in reliance on prediction. As far as we know, no experimental study has directly tested this hypothesis in the context of syntactic priming. Studies on lexical priming, however, have shown that drawing participants' attention to the relatedness between primes and targets, either explicitly (Holcomb, 1988) or implicitly by manipulating the proportion of semantically associated prime–target pairs (Lau et al., 2013), can increase the degree to which the prime affects the processing of the target. Similarly, Brothers et al. (2017) demonstrated that effects of lexical prediction during sentence processing were enhanced when participants were explicitly instructed to try and predict the last word of a passage, as well as when the overall validity of predictive cues across the experiment was increased. These findings suggest that top-down goals and strategies, both implicit and explicit, can modulate priming and prediction at the lexico-semantic level. Here we test whether explicit instructions to predict another speaker's description of a picture will increase the likelihood that the participant will subsequently produce the syntactic construction used by that speaker in their own descriptions. This study thus presents an initial attempt to extend the investigation of effects of top-down strategies on prediction to syntactic priming.

Prediction in L2 processing and structural priming

Apart from the participant-level and task-related factors examined in research with native speakers, additional factors such as language proficiency and native-speaker status have been proposed to modulate engagement in prediction during language processing among non-native and bilingual speakers (Kaan et al., 2010; Kaan, 2014; Grüter et al., 2017; Peters et al., 2018). For example, Grüter et al. (2017) proposed that non-native speakers have *Reduced Ability to Generate Expectations* (RAGE) during sentence and discourse processing. The RAGE proposal was motivated by findings showing reduced effects of prediction among L2 versus native (L1) speakers in sentence and discourse processing studies using both online (Grüter et al., 2012; Martin et al., 2013) and offline methodologies (Grüter et al., 2017). In the meantime, several studies have shown that L2 users are capable of engaging in predictive processing, sometimes to the same level as L1 users (e.g., Dijkgraaf et al., 2017). Others have reported smaller, delayed, or no effects of prediction among L2 users (e.g., Mitsugi & MacWhinney, 2016). Taken together, the current knowledge base suggests that L2 users can engage in prediction, but overall appear to do so to a lesser extent than native speakers. Why this is the case and what factors modulate L2 users' engagement in prediction, remain a matter of on-going investigation (for review, see Kaan, 2014; Kaan & Grüter, this volume).

What matters for present purposes is the descriptive generalization that, overall, L2 users appear to rely less on prediction during language processing than native speakers. In consequence, they should experience prediction errors less often, and thus encounter fewer opportunities to learn through the mechanisms proposed in accounts of error-based learning. Kaan and Chun (2018) explicitly appealed to this possibility in the context of a written structural priming study with Korean L2 learners of English targeting ditransitive constructions. Results showed cumulative adaptation effects in that the learners' likelihood to produce a construction increased the more often it had been encountered over the course of the experiment. Somewhat surprisingly, however, only weak and non-significant effects of immediate priming were observed. While acknowledging that such an explanation alone could not account for the full range of observations in their experiment, the authors suggested that "the L2 learners may not have actively predicted the elements after the verb in the prime sentence," and therefore "[u]nder error-based views of priming, the learners will not have received an error signal when encountering a DO prime, and hence will not have adjusted their representations in response to the prime structure" (Kaan & Chun, 2018, p. 239). This is fully consistent with both error-based learning accounts of structural priming and with what we know about the role of prediction in L2 processing. As an explanation of Kaan and Chun's findings, however, it remains speculative, as their experimental paradigm did not directly measure or manipulate participants' engagement in predictive processing.

Nevertheless, the possibility raised by Kaan and Chun (2018) has potentially important implications for applied linguistics and SLA, where the view of structural priming as implicit learning has given rise to the development of pedagogical applications that seek to integrate priming into curricular activities, in the hope that such activities will facilitate learning of targeted structures (e.g., McDonough & Chaikitmongkol, 2010). Outcomes of these studies have been variable (for review, see Jackson, 2018). We hypothesize that one reason why structural priming activities with L2 learners have shown variable outcomes is that learners partook in these activities with no or only limited engagement in prediction. In the present study, we test this hypothesis by presenting one group of L2 learners with a task that not only forced them to predict but also to assess whether their prediction was correct, while another group was presented with the same linguistic input in a task that required no prediction.

The role of L2 proficiency in predictive processing and structural priming

In a programmatic review of factors modulating L2 learners' engagement in prediction, Kaan (2014) included L2 proficiency as one such potential factor. Several studies have since investigated the role of proficiency in predictive processing directly. Somewhat surprisingly, despite a report of significant modulation of lexico-semantic prediction by proficiency in an early study by Chambers and Cooke (2009), more recent studies have reported no significant effects of proficiency on predictive processing at the level of lexico-semantics (Dijkgraaf et al., 2017; Ito et al., 2018), morpho-syntax (Hopp, 2015; Mitsugi, 2018), and reference resolution in discourse (Grüter et al., 2017; Kim & Grüter, 2021). Notably, the studies that found significant effects of proficiency all focused on grammatical gender (Dussias et al., 2013; Hopp, 2013; see also Hopp, this volume), and where reported, knowledge of gender assignment correlated strongly with overall L2 proficiency (Hopp, 2013). This makes it difficult to tease apart whether the modulation observed in these studies was due to knowledge of the specific linguistic property under investigation or L2 proficiency more generally. Thus, while proficiency remains a perhaps intuitive candidate for modulating engagement in prediction during L2 processing, the empirical evidence so far has been less conclusive than one might expect.

In the context of structural priming, there appears to be general consensus that *crosslinguistic* structural priming becomes stronger with increasing L2 proficiency (Van Gompel & Arai, 2018). This has been attributed to increasingly shared syntactic representations across languages as L2 representations become more abstract with increasing experience and proficiency (Hartsuiker & Bernolet, 2017). Fewer studies have examined the role of proficiency in structural priming within the L2. Lower proficiency L2 learners appear to show greater priming when lexical material is shared between primes and targets ("lexical boost," Pickering & Branigan, 1998; Kim & McDonough, 2008), an effect that may be due to explicit imitation strategies in the absence of more abstract syntactic representations among less proficient speakers (Hartsuiker & Bernolet, 2017). Experiments without shared lexical material have shown conflicting results: While Schoonbaert et al. (2007, reanalyzed in Hartsuiker & Bernolet, 2017) found increasing proficiency associated with smaller priming effects in a study with Dutch-English bilinguals targeting dative constructions in English, Bernolet et al. (2013) found greater priming effects for higher proficiency Dutch-English bilinguals on English genitive constructions. In an attempt to reconcile these findings, Hartsuiker and Bernolet (2017) suggested that in the absence of lexical overlap, we should generally see greater structural priming with increased proficiency as a result of the establishment of more abstract structural representations (analogous to the rationale for crosslinguistic structural priming), yet the explicit memory and imitation strategies that are assumed to account for the increased lexical boost effect in lower proficiency L2 speakers may also operate in the absence of lexical overlap depending on the nature of the construction involved.

In light of the limited and partially conflicting evidence from previous research on the role of proficiency in both predictive processing and structural priming in L2, we included an independent measure of proficiency in the present study with the primary goal of ensuring that proficiency did not present a confound in the critical comparison between the two experimental groups. In order to further explore potentially modulating effects of proficiency, we added proficiency scores as a continuous predictor to the model in a second, exploratory analysis step.

This study

The primary goal of this study is to test the hypothesis that arises from error-based accounts of learning, namely that increased engagement in prediction leads to increased learning. To this end, we conducted a written structural priming experiment with L1-Korean learners of English focusing on the dative alternation in English.

Previous work has shown that native speakers of (American) English frequently use both double-object (DO, *The girl fed the squirrel some nuts*) and prepositional dative (PO, *The girl fed some nuts to the squirrel*) constructions, with a potential overall preference for the former (Bock & Griffin, 2000; Jaeger & Snider, 2013). L2 learners of English, on the other hand, tend to have a strong preference for using POs. This has been shown for learners from a variety of L1 backgrounds (McDonough, 2006), and in particular for Korean learners of English (Kaan & Chun, 2018; Shin & Christianson, 2012). The fact that DOs constitute a dispreferred structure for Korean learners of English makes them an ideal target for the present study as priming should be facilitated by the inverse frequency effect. Conversely, priming would be difficult to show for POs under these circumstances since production of POs will likely be at or close to ceiling at baseline (see also Shin & Christianson, 2012). For this reason, only DOs are primed in this experiment (similar to Experiment 2 in McDonough, 2006). We will report effects of immediate priming, i.e., the likelihood of a DO produced on a target trial immediately following a (DO) prime, as well as longer-term adaptation, i.e., an increase in the production of DOs in an immediate post-test as compared to production at baseline.

Critically, participants will be assigned to one of two groups: the guessing-game (GG) condition, in which they will be given a task that forces them to predict and attend to prediction error, or the control condition (CC), in which participants will encounter the same linguistic materials but in a task that does not require prediction. This allows us to address the following primary research question:

RQ: Does forced engagement in prediction lead to greater effects of (1) immediate priming, and (2) longer-term adaptation as measured by change from baseline to an immediate post-test?

Based on error-driven models of learning and structural priming, we expect consistently greater priming in the GG than in the CC group. As a second and more exploratory question, we also examine the role of L2 proficiency on the size of priming effects in each phase and group. Given the inconsistency of previous findings on the role of proficiency, we make no predictions regarding effects of proficiency.

Methods

Participants

Thirty-five native Korean-speaking L2 learners of English from the University of Hawaiʻi student community, including students in short-term English language programs, participated in this study and were randomly assigned to one of two groups: guessing-game (GG, $N = 18$, 15 females) or control condition (CC, $N = 17$, 12 females).[1] In addition to the priming experiment, all participants completed an English cloze test (Brown, 1980) as well as a language background questionnaire, which included self-ratings of their English language skills. Table 1 presents a summary of participants' background information and results of pairwise comparisons. Due to the small sizes, analyses were conducted using non-parametric statistics (Mann-Whitney *U*, Spearman's rho). Pairwise comparisons indicated no significant differences between the two groups for any of the variables in Table 1. Cloze test scores and self-ratings were strongly correlated, $r_S(32) = .76$, $p < .001$.

Table 1. Participant demographics (means, standard deviations and ranges) and between-group comparisons

	Guessing Game (GG) group (*N* = 18)	**Control Condition (CC) group (*N* = 17)**	
Age	23.8 (*SD* = 5.1) (18–42)	22.9 (*SD* = 3.8) (19–34)	$U = 122$ $p = .31$
Age of first exposure to English	9.4 (*SD* = 4.0) (5–21)	9.1 (*SD* = 3.6) (3–18)	$U = 156$ $p = .92$
Length of stay in English speaking environment (in months)	16 (*SD* = 28) (0–108)	22 (*SD* = 36) (0–109)	$U = 147$ $p = .86$
Cloze test score (/50, acceptable-answer scoring)*	30.2 (*SD* = 10.0) (17–47)	26.8 (*SD* = 10.7) (9–45)	$U = 115.5$ $p = .33$
Self-rating of overall English language ability (0–10)	6.1 (*SD* = 1.7) (4–9)	5.5 (*SD* = 2.4) (2–10)	$U = 129.5$ $p = .44$

* One participant in the GG group did not complete the cloze test.

Materials

The overall structure of the priming task is summarized in Table 2; a complete list of all linguistic materials is available at https://osf.io/c3af4/. Twenty-two ditransitive verbs were selected from the materials used by Jaeger and Snider (2013, Appendix A). Only verbs that appeared on vocabulary lists in English textbooks used in Korean middle and high schools and in the vocabulary list for the Korean SAT test were included. Six ditransitive verbs were used for target trials in both the baseline and post-test phase, with the same verbs used in both task phases but

1. Data collection was cut short by the outbreak of the covid-19 pandemic, resulting in smaller than planned sample sizes. In view of previous production priming studies with similarly small *N*s (e.g., Kaschak et al., 2011), we proceeded with analysis, but acknowledge limited statistical power.

associated with different arguments (e.g., *give* was associated with *mother, girl, cake* in the baseline phase, and with *policeman, driver, ticket* in the post-test). Based on Jaeger and Snider's (2013) norming data, three ditransitive verbs in the baseline and post-test phase were biased towards a DO completion and three ditransitive verbs were biased towards a PO completion. The remaining 16 ditransitive verbs were used for prime and target trials in the priming phase, with 8 verbs used once each in primes and 8 verbs used once each in targets. No verbs or other lexical items were repeated between primes and targets. Based on Jaeger and Snider's (2013) norming data, five ditransitive verbs used in the prime sentences and five ditransitive verbs used in the target sentences were biased towards a DO completion and the remaining three ditransitive prime verbs and three ditransitive target verbs were biased towards a PO completion. Each sentence was paired with a colored clip-art image illustrating the event, with all arguments labelled and the verb printed below the image in its infinitive form (see Figure 1). In addition, 48 similarly formatted sentence-picture pairs were created using intransitive and simple transitive verbs. These constituted prime trials in the baseline and post-test phase, as well as prime and target trials in practice and filler items (see Table 2). Four filler prime-target pairs each were included in the baseline and post-test phase, and eight such pairs were included in the priming phase. Four experimental lists were created in which the order of items within each task phase (baseline, priming, post-test) was pseudorandomized. Thus, all participants saw the same items in each test phase but in different orders.

Table 2. Priming experiment: Structure and materials

Phase	**Experimental items**		**Fillers**	
	number and structure of prime-target pairs		**number and structure of prime-target pairs**	
Practice	0		2	prime: (in)transitive target: (in)transitive
Baseline	6	prime: (in)transitive target: ditransitive	4	prime: (in)transitive target: (in)transitive
Priming	8	prime: ditransitive: DO target: ditransitive	8	prime: (in)transitive target: (in)transitive
Posttest	6	prime: (in)transitive target: ditransitive	4	prime: (in)transitive target: (in)transitive

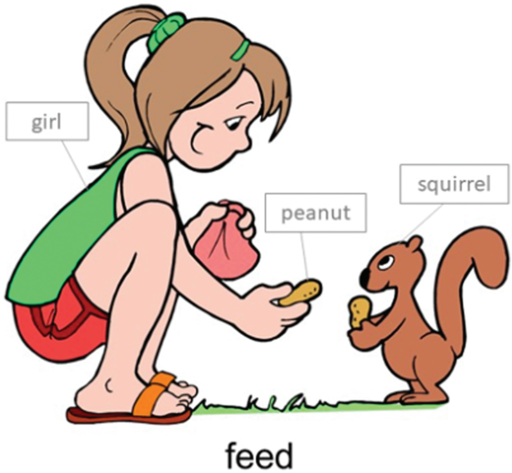

Figure 1. Example of visual stimuli

Procedure

All participants completed the language background questionnaire, followed by the priming experiment in either the GG or CC condition (described below). Immediately after the priming experiment, a semi-structured oral interview was conducted with the aim of evaluating the extent to which participants were (a) aware that the focus of the study was on the dative alternation, and (b) consciously trying to align their picture descriptions with those provided by the virtual partner. After the oral interview, participants completed the Brown cloze test, a 50-item written test that has shown high validity and reliability with comparable learner groups (Brown, 1980; Brown & Grüter, 2020). All testing was completed in a single session lasting approximately 60–90 minutes.

Guessing game (GG) condition

Participants in the GG condition were introduced to the experiment as follows: "In this activity, we would like to see how well you can GUESS how another person (Jessica) described pictures in English." They were then introduced to "Jessica," who was depicted and described to fit the stereotype of a native English-speaking American college student. Instructions then continued: "You will take turns guessing how Jessica described a picture, and describing pictures on your own. When the picture has a GREEN frame, your job is to guess how Jessica described the picture.

When the picture has a BLUE frame, your job is to describe the picture yourself. IMPORTANT: Please use ALL the words you see on the screen."

On prime trials in all three task phases, participants were presented with a labelled image, together with a picture of Jessica and the prompt "Write what you think Jessica wrote about this picture" (Figure 2, panel A). They then typed a sentence into a textbox. On the next screen (panel B), they were presented with Jessica's actual sentence – which always consisted of a DO construction in experimental items – together with the sentence participants had typed on the previous screen, and they were asked to indicate whether the two sentences were the same. This step was included to force participants to attend to potential differences between the predicted and the actual sentence, and thus to explicitly compute prediction error. This concluded the prime trial, and participants then proceeded to the next screen (target trial), where they were presented with an image as in Figure 1 and typed a sentence to describe it.

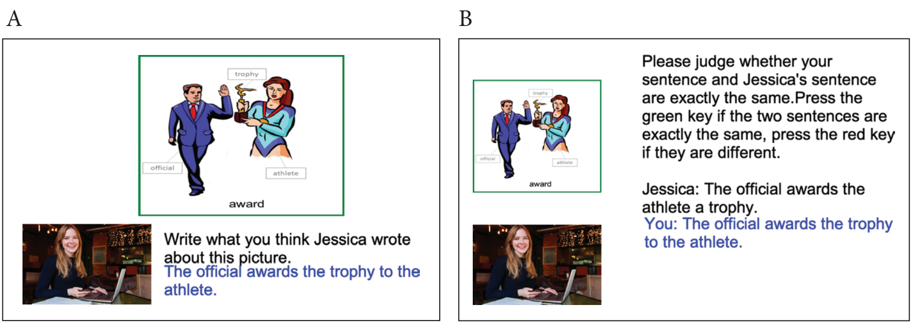

Figure 2. Prime trial in the guessing-game (GG) condition

Figure 3. Prime trial in the control condition (CC)

Control condition (CC)

Participants in the CC condition were introduced to the experiment as follows: “In this activity, you will learn and use English words, and you will practice putting them together to make sentences to describe pictures. You will read and copy sentences that another person (Jessica) wrote, and you will write your own sentences.” They were then introduced to Jessica in the same way as participants in the GG condition, and instructions continued: “You will take turns copying Jessica’s sentences, and describing pictures on your own. When the picture has a GREEN frame, your job is to read and copy Jessica’s sentence by typing it out again. When the picture has a BLUE frame, your job is to describe the picture yourself.”

For all task phases, prime trials in the CC condition consisted of a single screen in which participants were presented with a labelled image, together with a picture of Jessica, and Jessica’s description of the picture (Figure 3). Participants then re-typed Jessica’s sentence into a textbox. The copy–paste function was disabled so that participants were forced to re-type the sentence. They then proceeded to the next screen, consisting of target items in the same format as in the GG condition (Figure 1). The procedure in the control condition thus constitutes a typical production priming paradigm in the written mode. Production priming was chosen so that participants in both groups were engaged in the same basic activity, typing a sentence, with the critical difference that participants in the CC condition simply repeated Jessica’s sentence whereas those in the GG condition had to generate the sentence themselves in anticipation of what Jessica might have written.

In both conditions, participants completed two practice trials, which did not include ditransitives, before the beginning of the baseline phase, followed by the priming phase and the (immediate) post-test. As in other priming studies (e.g., Hartsuiker & Westenberg, 2000; Jackson & Ruf, 2018), transitions between these three task phases were not signaled to the participant, such that participants viewed the priming activity as one continuous task. The experiment was conducted in PsychoPy 2.0 (Peirce et al., 2019).

Data annotation and analysis

Sentences produced in target trials were annotated as “DO” when they contained a recipient NP followed by a theme NP and no preposition (e.g., *The mother gives the girl a cake,* 28.1% of all responses), as “PO_to” when it contained a theme NP, followed by *to* and a recipient NP (e.g., *Mom gives a cake to the girl*, 54.2%), and as “PO_otherPrep” when the recipient was preceded by a preposition other than *to* (e.g., *The mom gives the cake for girl*, 8.7%). These three sentence types together accounted for 91.0% of all target sentences produced (CC: 91.2%, GG: 90.8%). Sentences not fitting these criteria were excluded from further analysis (e.g., *The bride text message and send to groom*). "PO_to" and "PO_otherPrep" responses were collapsed for further analysis into a single category "PO". Non-target inflectional marking (tense, number) and article usage was disregarded for coding purposes. Sentences produced by participants in the GG condition on prime trials were annotated following the same criteria. DO and PO responses accounted for 85.4% of these data.

All analyses are based on data points consisting of DO and PO only. Mixed-effect logistic regression was employed to predict the likelihood of a DO sentence across task phase (baseline, priming, posttest) and group (GG, CC). Since the two groups did not differ by proficiency (see Table 1) and to avoid overfitting, proficiency was not included as a factor in the primary models. In order to explore the role of proficiency, we added cloze test scores to the model in a second step. Cloze test scores, rather than participants' self-ratings, were included as the measure of proficiency given the established reliability and validity of this cloze test with similar participant samples (Brown & Grüter, 2020). All analyses were conducted in R 3.6.0 (R Core Team, 2019) using the lmerTest package (Kuznetsova et al., 2017).

Results

Figure 4 presents the proportion of DOs produced by participants in both groups across the three task phases (baseline, priming, post-test). The production of DOs at baseline was low in both groups (M_{CC} = .07, SD = .13; M_{GG} = .12, SD = .22), with

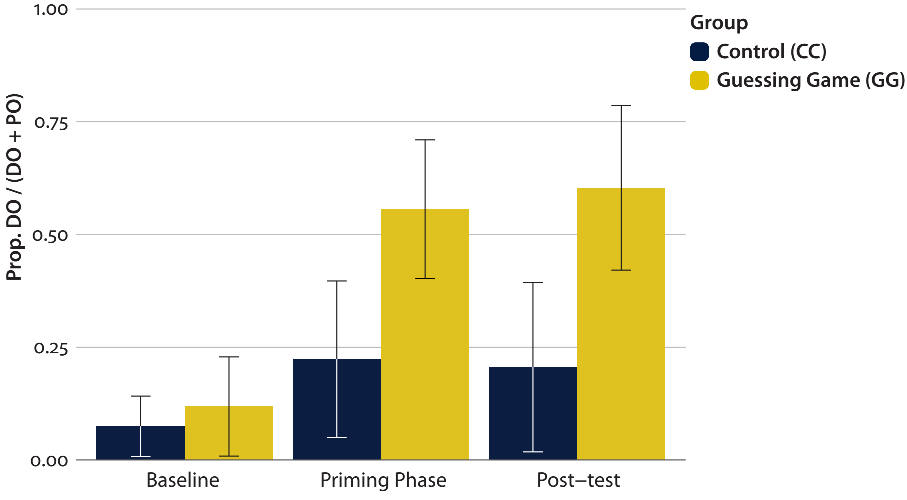

Figure 4. Proportion of DOs out of all DO and PO utterances by group and task phase. Error bars indicate 95% confidence intervals on means by participants.

only 5 (of 17) and 6 (of 18) participants in the CC and GG groups, respectively, producing DOs at all. The proportion of DOs increased in both groups in the priming phase (M_{CC} = .22, *SD* = .34; M_{GG} = .56, *SD* = .31), with 7 CC participants and 16 GG participants producing DOs. These proportions remained similar in the post-test (M_{CC} = .21, *SD* = .37; M_{GG} = .60, *SD* = .37), with 7 CC and 16 GG participants producing DOs.

The likelihood of producing a DO was assessed in a mixed-effects logistic regression model with group (centered and contrast-coded, −.5 = CC, .5 = GG), task phase (dummy-coded, reference level: baseline) and their interaction included as fixed effects, and participants and items as random effects.[2] The output from this model is presented in Table 3.

Table 3. Model statement and summary of fixed effects in mixed effects logistic regression Formula: DO ~ Group * TaskPhase + (1 | Subject) + (1 | Item)

Fixed effects:	**Estimate**	***SE***	***z***	***p***
(Intercept)	−3.47	0.53	−6.52	< 0.001
Group	1.06	0.88	1.21	0.23
TaskPhase (baseline-priming)	2.59	0.49	5.32	< 0.001
TaskPhase (baseline-posttest)	2.60	0.51	5.09	< 0.001
Group : TaskPhase (baseline-priming)	1.32	0.69	1.90	0.06
Group : TaskPhase (baseline-posttest)	1.62	0.72	2.24	0.03

Across both groups, significant increases in the likelihood of DOs were observed in the priming vs. baseline phase ($b = 2.59, p < .001$), as well as in the post-test vs. baseline ($b = 2.60, p < .001$). These effects were qualified by interactions with group, marginal for the increase in the priming phase, $b = 1.32, p = .06$, and fully significant for the increase from baseline to post-test, $b = 1.62, p = .03$. To further explore the interactions, separate models were fit to the data from each group. For the GG group, the production of DOs increased significantly compared to baseline in both the priming phase ($b = 3.14, p < .001$) and the post-test ($b = 3.28, p < .001$). The same was found for the CC group, albeit with smaller effect sizes (baseline-priming: $b = 1.81$, $p < .001$; baseline-posttest: $b = 1.65, p = .003$).[3] In sum, these results show priming effects in both groups, for both immediate priming and longer term adaptation as

2. The maximal random effects structure justified by the design should include random participant slopes for task phase and random item slopes for group. The inclusion of these slopes was attempted, but led to boundary singularity errors. The largest model that fully converged included random intercepts only.

3. The model for the CC group data did not converge with random intercepts for both participants and items. Items were thus removed from the random effects structure of this model.

measured by performance on the post-test. Critically, effects were larger in the GG than in the CC group, with the between-group difference approaching significance in the priming phase and reaching full significance in the post-test.

Production of predicted primes in the GG group

In order to examine the extent to which GG participants were successful in predicting that Jessica would produce DOs on ditransitive prime items during the priming phase, we analyzed their responses on prime items, i.e., the sentences they guessed that Jessica would produce to describe ditransitive events. Out of the 8 ditransitive primes, GG participants produced a mean of 3.0 DOs (*SD* = 2.0); no participant produced DOs on more than 6 out of 8 trials (range: 0-6). For comparison with the proportion of DOs produced on target items (Figure 4), we also calculated the mean proportion of DOs out of DO and PO responses only: *M* = .44 (*SD* = .29). These results indicate that despite the explicit task of guessing what Jessica would say and Jessica's consistent production of DOs, participants were far from consistent in producing DOs when predicting what Jessica would say.

Proficiency

To explore whether L2 proficiency modulated the size of priming effects in this study, we added cloze test scores (centered) as an additional continuous fixed effect to the model reported in Table 3. This model, including all three predictors of interest – group, task phase, and proficiency – and their interactions, failed to converge. We therefore proceeded to conduct two separate models, one including only group, proficiency and their interaction (but not task phase, Model 1), the other including only task phase, proficiency and their interaction (but not group, Model 2).

Model 1 (DO ~ Group * scale(Proficiency) + (1 | Subject) + (1 | Item)) yielded significant main effects of group (b = 2.02, SE = .67, z = 3.01, p = .003) and proficiency (b = .87, SE = .34, z = 2.58, p = .01), but no interaction (b = .32, SE = .66, z = .48, p = .63). The positive estimate for proficiency indicates that more proficient participants were more likely to produce DOs overall. The lack of an interaction indicates that this was the case to the same extent for participants in both groups. The main effect of group shows that the two groups differed in their production of DOs across the experiment even when variability associated with proficiency was accounted for, thus providing further assurance that proficiency was not a confound in our primary analysis.

Model 2 (DO ~ TaskPhase * scale(Proficiency) + (1 | Subject) + (1 | Item)) showed the main effects of task phase observed in the primary model (baseline-priming:

$b = 4.12$, $SE = .78$, $z = 5.30$, $p < .001$; baseline-posttest: $b = 4.13$, $SE = .80$, $z = 5.19$, $p < .001$), as well as a main effect of proficiency ($b = 2.80$, $SE = .67$, $z = 4.17$, $p < .001$), which was qualified by interactions with task phase (Proficiency*baseline-priming: $b = -1.79$, $SE = .56$, $z = -3.19$, $p = .001$; Proficiency*baseline-posttest: $b = -2.09$, $SE = .58$, $z = -3.63$, $p < .001$). To better understand the interactions between proficiency and task phase, we visualized the marginal effects of these interaction terms using the *interactions* package in R (Long, 2019). As Figure 5 illustrates, increased production of DOs from baseline to priming and post-test phases is greatest among participants in the middle of the proficiency distribution. For those at the lower end, increases are small, as they produced few or no DOs throughout the experiment. At the higher end of the spectrum, participants produced DOs more consistently at baseline and DOs were also produced in later phases, but the difference between baseline and the subsequent two task phases was not as large as for participants in the middle of the proficiency range, who produced few or no DOs at baseline, and increased their production drastically over the course of the experiment.

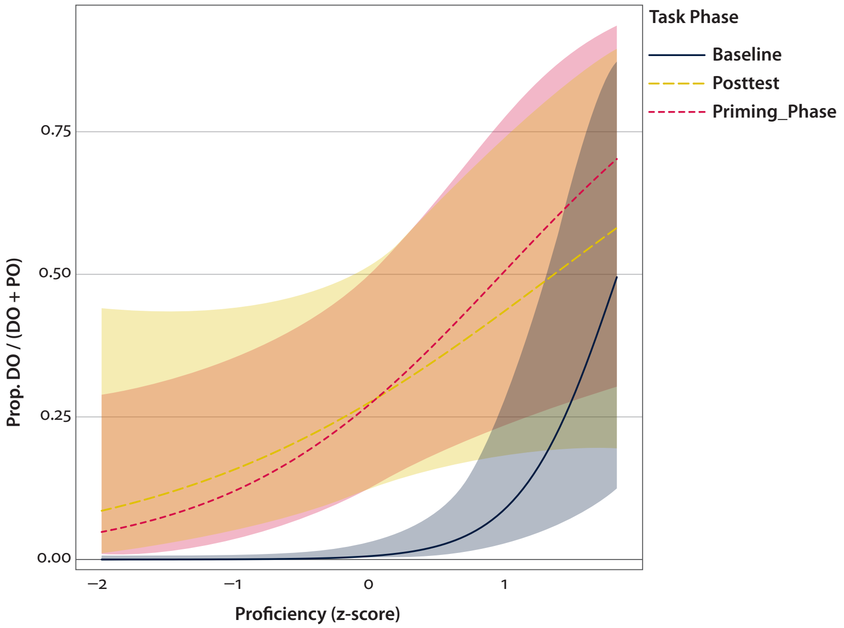

Figure 5. Visualization of marginal interaction terms using the *interact_plot* function in R; data from both groups combined. Error bands show 95% CIs.

Despite the absence of an interaction between proficiency and group in Model 1, we wanted to explore to what extent the interaction between proficiency and task phase observed in Model 2 obtained within each group. We therefore fit Model 2 to the data from each group separately. For both groups, the output mirrored the overall pattern, with significant main effects of task phase, proficiency, and negative interactions between the two (CC: Proficiency*baseline-priming: $b = -2.06$, $SE = .88$, $z = -2.33$, $p = .02$, Proficiency*baseline-posttest: $b = -2.48$, $SE = .93$, $z = -2.67$, $p = .008$; GG: Proficiency*baseline-priming: $b = -1.67$, $SE = .70$, $z = -2.37$, $p = .02$, Proficiency*baseline-posttest: $b = -1.94$, $SE = .73$, $z = -2.66$, $p = .008$). These results provide further indication that the role of proficiency was similar in the two groups.

Exit interview

The goal of the exit interview was to gauge whether participants may have employed explicit strategies, a possibility that seemed particularly relevant since the dative alternation is explicitly taught in EFL classrooms in Korea. To this end, we annotated and analyzed participants' responses to (1) the yes-no question "Do you think the way Jessica described the pictures influenced how YOU described them?," and (2) the follow-up question "In what way?" to a yes-response on the previous question. Responses were scored by two independent annotators, scoring (1) for 1 = yes, 0 = no, and (2) for 1 = mention of dative alternation in some form, 0 = mention of another linguistic property (e.g., article use), NA = no answer. Interrater agreement was high (34/35 and 33/35 respectively), and disagreements were resolved through discussion.

The majority of CC participants (13/17) and all GG participants (18/18) answered yes to the first question, indicating that most participants actively tried to align their linguistic choices with those of a virtual partner. Of the 13 CC participants who indicated that they actively sought to align with Jessica, 8 explicitly mentioned that they tried to do so by using more DOs. Other properties mentioned were usage of articles and verb tense. In the GG group, all but one participant (17/18) explicitly mentioned the dative alternation. These observations suggest that the majority of participants in both groups actively tried to align their responses with the linguistic models provided by a native speaker, but those in the GG condition were overall more likely to focus on the dative alternation.

Discussion

Motivated by previous work on error-driven learning and implicit learning accounts of structural priming, we set out to test whether forced engagement in prediction would lead to greater effects of (1) immediate priming, and (2) longer-term adaptation as measured by change from baseline to an immediate post-test. Korean L2 learners of English were presented with double-object primes from a virtual partner, Jessica. Those in the GG condition had to predict Jessica's descriptions and then compare their predictions to Jessica's actual descriptions; those in the CC condition merely had to retype Jessica's descriptions, following a standard production priming procedure. Both groups exhibited immediate priming, with significantly more DOs produced in the priming phase following DO primes than at baseline, as well as longer term adaptation as reflected by continuedly increased production of DOs at post-test. Critically, the increase in DO production at post-test was significantly greater in the GG than in the CC group. The persistence of priming beyond immediately adjacent trials is consistent with error-driven accounts of structural priming (Bock & Griffin, 2000; Dell & Chang, 2014) and suggests that the effects observed in both groups are unlikely to be driven by explicit memory or imitation strategies alone (Hartsuiker & Bernolet, 2017; Shin & Christianson, 2012). We note, however, that conclusions about longer-term adaptation and learning based on performance on an immediate post-test must be drawn with caution. While this was not logistically feasible in the present study, evidence from a delayed post-test would be desirable to better understand the permanence of what we interpret as learning effects and the observed difference between the forced-prediction (GG) and no-prediction (CC) treatments. Nevertheless, the observation that the increase in the use of DOs was substantially greater in the GG compared to the CC group, in both the priming phase and the immediate post-test, and both when measured in terms of mean proportion DOs produced and proportion of participants producing DOs, strongly suggests that being forced to predict the prime sentence facilitated L2 learners' use of DOs in their own subsequent productions.

Exploratory analyses including L2 proficiency showed no difference between the two groups, but revealed an overall pattern that may help reconcile contradictory findings from previous within-L2 priming studies. Our analyses showed consistent interactions between proficiency and the size of priming effects, both from baseline to the priming phase and from baseline to post-test. The visualization of these interactions revealed a non-linear relation, with less priming among the lowest and highest proficient participants, and greater priming among those in the middle of the distribution. The absence of priming among low-proficiency speakers is consistent with the proposal that priming requires some form of abstract representation of the structure in question (Hartsuiker & Bernolet, 2017; McDonough & Fulga, 2015). These low-proficiency participants may not yet have had a sufficiently stable representation of DOs to benefit from error-driven learning. Reduced priming among our highest proficiency learners, on the other hand, is consistent with surprisal-based accounts, in that Figure 5 indicated that this reduced effect was driven by greater likelihood to produce DOs at baseline. This suggests that DOs caused less surprisal among these participants, and thus less adaptation as per the inverse frequency effect. The observation that priming was most effective among learners in the middle range of the proficiency spectrum suggests that there might be a "sweet spot" in development when the benefits of error-driven learning are greatest. We hypothesize that this is when an abstract representation has begun to be established but is not yet stable enough to support production in the absence of activation through priming.

We included an exit interview in this study to gauge whether participants engaged in explicit processes during the priming experiment. It is generally assumed that participants in priming experiments are unaware of their adaptation to primed structures and the nature of the structures under investigation (Bock & Griffin, 2000). There is strong evidence that priming occurs under conditions where explicit strategies are not available (Ferreira et al., 2008), yet few priming studies with healthy L1 or L2 adults report to what extent participants were aware of the purpose of the experiment (for exceptions, see Jackson & Ruf, 2018; Myslín & Levy, 2015). Results from the exit interview show that many participants in both groups were (a) aware of the construction under investigation, and (b) consciously tried to align their productions with those encountered on primes. This observation suggests that the assumption that priming is fully unconscious may not always be warranted, especially in studies with L2 learners. Thus, conclusions about *implicit* learning should be drawn with caution, and future studies addressing these issues – both with L1 and L2 users – may benefit from including a measure of participants' explicit awareness.

Responses on the exit interview indicated that while the majority of participants in both groups consciously tried to align with the virtual partner, awareness of that partners' use of DO constructions was reported by more participants in the GG (17/18) than in the CC (8/17) group. Given the small sample sizes in the present study, we refrain from further statistical comparison of subgroups of participants who did and did not report awareness of DOs. We note, however, that such comparisons would be valuable in a larger study to better understand to what extent the increased adaptation effects observed in the GG group derive from the forced-prediction treatment leading to increased attention and awareness more generally, rather than the computation of prediction error as we hypothesized here. Yet to the extent that awareness is related to prediction violation (Lupyan & Clark, 2015), dissociating the two may turn out to be less than straightforward.

More generally, it is clear that additional work is needed to better understand which aspects of the guessing-game manipulation led to the increased priming effects observed in this study. With the goal of involving participants in the control condition in as similar a task as possible, we had decided on a production priming paradigm in which CC participants had to retype the prime sentence. Thus, all participants *produced* a ditransitive sentence on prime trials; however, only GG participants actively *generated* one and explicitly compared their own utterance against a model. We therefore cannot exclude the possibility that the observed between-group difference reflects better learning due to other mechanisms involved in generating vs repeating a sentence, rather than due to the computation of prediction error. Recent work by Potts et al. (2019) on novel vocabulary learning has shown better learning success when participants had to guess the meaning of a novel word first than when they were presented with form-meaning pairs from the start. Importantly, this advantage emerged only in contexts where there was an information gap that participants had a desire to fill. This led the authors to propose that "the act of generating a response to an unfamiliar cue, where the correct response is not yet known, stimulates a desire to close an information gap, leading to enhanced motivation to encode corrective feedback" (Potts et al., 2019, p. 1039; see also Gambi, this volume). It therefore remains possible that the increased priming effects observed in the GG group in this study were due in part to increased motivation to close an information gap present only in the GG condition. Future work is needed to tease apart what specific aspects of the GG treatment beside the computation of prediction error contributed to the greater effects of priming and adaptation that were observed in this study. This will enhance not only our understanding of the role of prediction in L2 processing and learning, but ultimately our ability to harness these benefits to support L2 learning.

In conclusion, the small-scale study reported here presents a first step towards the investigation of how explicit manipulation of L2 learners' engagement in prediction may affect learning in the form of structural priming and adaptation. Findings lend some support to the hypothesis that forced prediction can increase both immediate priming and longer term adaptation, suggesting that the explicit computation of prediction error may benefit L2 learning in the sense of increased likelihood to produce a previously dispreferred construction. However, the present findings do not allow us to fully tease apart to what extent the differences observed between the two groups in this study are attributable to the computation of prediction error by participants in the guessing game condition, and to what extent other factors related to the guessing-game manipulation, such as increased awareness and motivation to close an information gap, contributed to these different outcomes. Separating these explanations at both empirical and conceptual levels is a challenge that we must leave for future work to address.

Acknowledgments

We are grateful to Hyunwoo Kim for assistance with the creation of experimental materials for this study. We thank Holger Hopp, Edith Kaan and two anonymous reviewers for their very helpful feedback on an earlier version of this paper.

References

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Science*, 11(7), 280–289. https://doi.org/10.1016/j.tics.2007.05.005

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, 127(3), 287–306. https://doi.org/10.1016/j.cognition.2013.02.005

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177–192. https://doi.org/10.1037/0096-3445.129.2.177

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Goals and strategies influence lexical prediction during sentence comprehension. *Journal of Memory and Language*, 93, 203–216. https://doi.org/10.1016/j.jml.2016.10.002

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, 64(3), 311–317. https://doi.org/10.1111/j.1540-4781.1980.tb05198.x

Brown, J. D., & Grüter, T. (2020). The same cloze for all occasions? Using the Brown (1980) cloze test for measuring proficiency in SLA research. *International Review of Applied Linguistics in Language Teaching*. Advance online publication. https://doi.org/10.1515/iral-2019-0026

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35(4), 1029–1040. https://doi.org/10.1037/a0015901

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1634), Article 20120394. https://doi.org/10.1098/rstb.2012.0394

Dijkgraaf, A., Hartsuiker, R., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20(5), 917–930. https://doi.org/10.1017/S1366728916000547

Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387. https://doi.org/10.1017/S0272263112000915

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, 27(6), 443–448. https://doi.org/10.1177/0963721418794491

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, 69(3), 165–182. https://doi.org/10.1016/j.jml.2013.06.001

Ferreira, V. S., Bock, J. K., Wilson, M. P., & Cohen, N. J. (2008). Memory for syntax despite amnesia. *Psychological Science*, 19(9), 940–946. https://doi.org/10.1111/j.1467-9280.2008.02180.x

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191–215. https://doi.org/10.1177/0267658312437990

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism*, 7(2), 199–229. https://doi.org/10.1075/lab.15011.gru

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, 20(2), 219–234. https://doi.org/10.1017/S1366728915000164

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, 75(2), B27–B39. https://doi.org/10.1016/S0010-0277(99)00080-3

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, 35(1), 66–85. https://doi.org/10.1016/0093-934X(88)90101-0

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, 29(1), 33–56. https://doi.org/10.1177/0267658312461803

Hopp, H. (2015). Semantics and morphosyntax in L2 predictive sentence processing. *International Review of Applied Linguistics in Language Teaching*, 53(3), 277–306. https://doi.org/10.1515/iral-2015-0014

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. https://doi.org/10.1016/j.brainres.2015.02.014

Huettig, F., & Pickering, M. J. (2019). Literacy advantages beyond reading: Prediction of spoken language. *Trends in Cognitive Science*, 23(6), 464–475. https://doi.org/10.1016/j.tics.2019.03.008

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017). On predicting form and meaning in a second language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(4), 635–652. https://doi.org/10.1037/xlm0000315

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. *Bilingualism, Language and Cognition*, 21(2), 251–264. https://doi.org/10.1017/S1366728917000050

Jackson, C. N. (2018). Second language structural priming: A critical review and directions for future research. *Second Language Research*, 34(4), 539–552. https://doi.org/10.1177/0267658317746207

Jackson, C. N., & Hopp, H. (2020). Prediction error and implicit learning in L1 and L2 syntactic priming. *International Journal of Bilingualism*, 24(5–6), 895–911. https://doi.org/10.1177/1367006920902851

Jackson, C. N., & Ruf, H. T. (2018). The importance of prime repetition among intermediate-level second language learners. *Studies in Second Language Acquisition*, 40(3), 677–692. https://doi.org/10.1017/S0272263117000365

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, 4(2), 257–282. https://doi.org/10.1075/lab.4.2.05kaa

Kaan, E., & Chun, E. (2018). Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, 21(2), 228–242. https://doi.org/10.1017/S1366728916001231

Kaan, E., Dallas, A. C., & Wijnen, F. (2010). Syntactic predictions in second-language sentence processing. In C. J.-W. Zwart & M. de Vries (Eds.), *Structure preserved: Studies in syntax for Jan Koster* (pp. 207–213). John Benjamins. https://doi.org/10.1075/la.164.23kaa

Kaschak, M. P., Kutta, T. J., & Coyle, J. M. (2014). Long and short term cumulative structural priming effects. *Language, Cognition and Neuroscience*, 29(6), 728–743. https://doi.org/10.1080/01690965.2011.641387

Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, 39(3), 381–388. https://doi.org/10.3758/s13421-010-0042-3

Kim, H., & Grüter, T. (2021). Predictive processing of implicit causality in a second language: a visual-world eye-tracking study. *Studies in Second Language Acquisition* 43(1), 133–154. https://doi.org/10.1017/S0272263120000443

Kim, Y., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics*, 29(1), 149–154. https://doi.org/10.1093/applin/amn004

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 effects of prediction from association in single-word contexts. *Journal of Cognitive Neuroscience*, 25(3), 484–502. https://doi.org/10.1162/jocn_a_00328

Long, J. A. (2019). *interactions: Comprehensive, user-friendly toolkit for probing interactions* (R package version 1.1.0) [Computer software]. https://cran.r-project.org/package=interactions

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science*, 24(4), 279–284. https://doi.org/10.1177/0963721415570732

Martin, C. D., Thierry, G., Kuipers, J., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, 69(4), 574–588. https://doi.org/10.1016/j.jml.2013.08.001

McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition*, 28(2), 179–207. https://doi.org/10.1017/S0272263106060098

McDonough, K., & Chaikitmongkol, W. (2010). Collaborative syntactic priming activities and EFL learners' production of wh-questions. *The Canadian Modern Language Review*, 66(6), 817–42. https://doi.org/10.3138/cmlr.66.6.817

McDonough, K., & Fulga, A. (2015). The detection and primed production of novel constructions. *Language Learning*, 65(2), 326–357. https://doi.org/10.1111/lang.12103

Meijs, E. L., Slagter, H. A., de Lange, F. P., & van Gaal, S. (2018). Dynamic interactions between top–down expectations and conscious awareness. *The Journal of Neuroscience*, 38(9), 2318–2327. https://doi.org/10.1523/JNEUROSCI.1952-17.2017

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5(1), 1–10. https://doi.org/10.16910/jemr.5.1.3

Mitsugi, S. (2018). Generating predictions based on semantic categories in a second language: A case of numeral classifiers in Japanese. *International Review of Applied Linguistics in Language Teaching*, 58(3), 323–349. https://doi.org/10.1515/iral-2017-0118

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, 19(1), 19–35. https://doi.org/10.1017/S1366728914000881

Myslín, M., & Levy, R. (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, 147, 29–56. https://doi.org/10.1016/j.cognition.2015.10.021

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peters, R. E., Grüter, T., & Borovsky, A. (2018). Vocabulary size and native speaker self-identification influence flexibility in linguistic prediction among adult bilinguals. *Applied Psycholinguistics*, 39(6), 1439–1469. https://doi.org/10.1017/S0142716418000383

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134(3), 427–459. https://doi.org/10.1037/0033-2909.134.3.427

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, 144(10), 1002–1044. https://doi.org/10.1037/bul0000158

Potts, R., Davies, G., & Shanks, D. R. (2019). The benefit of generating errors during learning: what is the locus of the effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(6), 1023–1041. https://doi.org/10.1037/xlm0000637

R Core Team. (2019). *R: A language and environment for statistical computing* [Computer Software]. R Foundation for Statistical Computing. https://www.R-project.org

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In W. E. Prokasy & A. H. Black (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56(2), 153–171. https://doi.org/10.1016/j.jml.2006.10.002

Shin, J. A., & Christianson, K. (2012). Syntactic processing in Korean-English: Evidence from cross-linguistic structural priming. *Cognition*, 112(1), 175–180. https://doi.org/10.1016/j.cognition.2009.03.011

Sorensen, D. W., & Bailey, K. G. D. (2007). The world is too much: Effects of array size on the link between language comprehension and eye movements. *Visual Cognition*, 15(1), 112–115. https://doi.org/10.1080/13506280600975486

Van Gompel, R. P., & Arai, M. (2018). Structural priming in bilinguals. *Bilingualism: Language and Cognition*, 21(3), 448–455. https://doi.org/10.1017/S1366728917000542

Wlotko, E. W., Federmeier, K. D., & Kutas, M. (2012). To predict or not to predict: Age-related differences in the use of sentential context. *Psychology and Aging*, 27(4), 975–988. https://doi.org/10.1037/a0029206

Index

There is ample evidence that language users, including second-language (L2) users, can predict upcoming information during listening and reading. Yet it is still unclear when, how, and why language users engage in prediction, and what the relation is between prediction and learning. This volume presents a collection of current research, insights, and directions regarding the role of prediction in L2 processing and learning. The contributions in this volume specifically address how different (L1-based) theoretical models of prediction apply to or may be expanded to account for L2 processing, report new insights on factors (linguistic, cognitive, social) that modulate L2 users' engagement in prediction, and discuss the functions that prediction may or may not serve in L2 processing and learning. Taken together, this volume illustrates various fruitful approaches to investigating and accounting for differences in predictive processing within and across individuals, as well as across populations.

ISBN 978 90 272 0970 2

9 789027 209702

John Benjamins Publishing Company