Linguistic Foundations of Narration in Spoken and Sign Languages

Edited by

Annika Hübl

Markus Steinbach

Linguistic Foundations of Narration in Spoken and Sign Languages

Linguistik Aktuell/Linguistics Today (LA)

Linguistik Aktuell/Linguistics Today (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/la

Volume 247

Linguistic Foundations of Narration in Spoken and Sign Languages
Edited by Annika Hübl and Markus Steinbach

Linguistic Foundations of Narration in Spoken and Sign Languages

Edited by

Annika Hübl
Markus Steinbach
University of Göttingen

John Benjamins Publishing Company

Amsterdam / Philadelphia

Table of contents

List of contributors

Gemma Barberà
CNRS/Paris 8
UMR 7023 Structures Formelles
du Langage
CNRS UPS-Pouchet
59/61, rue Pouchet
75017 Paris
France
Email: gemma.barbera@upf.edu

Jana Bressem
Technische Universität Chemnitz
Institut für Germanistik und
Kommunikation
Thüringer Weg 11
09126 Chemnitz
Germany
Email: jana.bressem@phil.tu-chemnitz.de

Mary Carroll
Ruprecht-Karls-Universität Heidelberg
Institut für Deutsch als
Fremdsprachenphilologie
Plöck 55
69117 Heidelberg
Germany
Email: carroll@idf.uni-heidelberg.de

Annika Herrmann
Universität Hamburg
Institut für Deutsche Gebärdensprache
Binderstraße 34
20146 Hamburg
Germany
Email: annika.herrmann@uni-hamburg.de

Anke Holler
Georg-August-Universität Göttingen
Seminar für Deutsche Philologie
Käte-Hamburger-Weg 3
37073 Göttingen
Germany
Email: anke.holler@phil.uni-goettingen.de

Annika Hübl
Georg-August-Universität Göttingen
Courant Forschungszentrum
„Textstrukturen"
Nikolausberger Weg 23
37073 Göttingen
Germany
Email: annika.huebl@phil.
uni-goettingen.de

Franziska Köder
University of Oslo
Center for the Study of Mind in Nature
Blindernveien 31
0313 Oslo
Norway
Email: franziska.koder@ifikk.uio.no

Silva H. Ladewig
Europa-Universität Viadrina
Kulturwissenschaftliche Fakultät
Große Scharrnstraße 59
15230 Frankfurt (Oder)
Germany
Email: ladewig@europa-uni.de

Evie Malaia
Purdue University
Linguistics, College of Liberal Arts
Speech, Language & Hearing Sciences,
College of Health and Human Sciences
Sign Language Linguistics Laboratory,
Lyles-Porter Hall
West Lafayette, IN 47907
United States of America
Email: emalaya@purdue.edu.

Cornelia Müller
Europa-Universität Viadrina
Kulturwissenschaftliche Fakultät
Große Scharrnstraße 59
15230 Frankfurt (Oder)
Germany
Email: cmueller@europa-uni.de

Nina-Kristin Pendzich
Georg-August-Universität Göttingen
Seminar für Deutsche Philologie
Käte-Hamburger-Weg 3
37073 Göttingen
Germany
Email: nina-kristin.pendzich@phil.uni-
goettingen.de

Svetlana Petrova
Bergische Universität Wuppertal
Fakultät für Geistes- und
Kulturwissenschaften
Gaußstraße 20
42119 Wuppertal
Germany
Email: petrova@uni-wuppertal.de

Josep Quer
ICREA-Universitat Pompeu Fabra
Departament de Traducció i Ciències
del Llenguatge
Roc Boronat 138
08018 Barcelona
Spain
Email: josep.quer@upf.edu

Susanna Salem
Georg-August-Universität Göttingen
Courant Forschungszentrum
„Textstrukturen"
Nikolausberger Weg 23
37073 Göttingen
Germany
Email: susanna.salem@zentr.uni-
goettingen.de

Augustin Speyer
Universität des Saarlandes
Fakultät P, FR Germanistik,
Campus A 2.2
66123 Saarbrücken
Germany
Email: a.speyer@mx.uni-saarland.de

Markus Steinbach
Georg-August-Universität Göttingen
Seminar für Deutsche Philologie
Käte-Hamburger-Weg 3
37073 Göttingen
Germany
Email: msteinb@gwdg.de

Christiane von Stutterheim
Ruprecht-Karls-Universität Heidelberg
Institut für Deutsch als
Fremdsprachenphilologie
Plöck 55
69117 Heidelberg
Germany
Email: stutterheim@idf.uni-heidelberg.de

Naoko Tomita
Ruprecht-Karls-Universität Heidelberg
Institut für Deutsch als
Fremdsprachenphilologie
Plöck 55
69117 Heidelberg
Germany
Email: naoko-tomita@nifty.com

Thomas Weskott
Georg-August-Universität Göttingen
Seminar für Deutsche Philologie
Käte-Hamburger-Weg 3
37073 Göttingen
Germany
Email: thomas.weskott@gmail.com

Ronnie B. Wilbur
Purdue University
Linguistics, College of Liberal Arts
Speech, Language & Hearing Sciences,
College of Health and Human Sciences
Lyles-Porter Hall
West Lafayette, IN 47907
United States of America
Email: wilbur@purdue.edu

Sonja Zeman
Otto-Friedrich-Universität Bamberg
Institut für Germanistik
Hornthalstraße 2
96047 Bamberg
Germany
Email: sonja.zeman@uni-bamberg.de

List of abbreviations

ASL	American Sign Language
A-RS	action role shift
BSL	British Sign Language
CA	constructed action
CB	backward-looking center
CF	forward-looking center
CP	preferred center
CT	Centering Theory
DEMPR	demonstrative pronoun
DGS	German Sign Language (*Deutsche Gebärdensprache*)
DP	determiner phrase
DRT	Discourse Representation Theory
ENHG	Early New High German
FID	Free Indirect Discourse
HKSL	Hong Kong Sign Language
HZJ	Croatian Sign Language (*Hrvatskom Znakovnom Jeziku*)
HP	historical present
Israeli SL	Israeli Sign Language
L1	first language
L2	second language
LSC	Catalan Sign Language (*Llengua de signes catalana*)
MHG	Middle High German
NP	noun phrase
OHG	Old High German
PERPR	personal pronoun
POV	point of view predicate
PP	prepositional phrase
Q-RS	quotation role shift
QUD	question under discussion
SASL	South African Sign Language
V2	verb-second (word order)

Notational conventions for examples from sign languages

SIGN	sign
IX$_{\#}$	INDEX, digit indicates referential loci in signing space
$_{\#}$VERB$_{\#}$	agreement verb, digits indicate referential loci of subject and object in signing space
CL	classifier
+++	reduplication of a sign (e.g. in order to express plural)
:	pause/prosodic break
SIGN $\overline{\text{SIGN SIGN}}$	simulateneous use of nonmanual markers and signs, small letters on the line refer to the form or to the function of the nonmanual marker

Approaching narration across modalities: Topics, methods, perspectives

Annika Hübl and Markus Steinbach
University of Göttingen

1. Introduction

For a considerable time, linguists have not only investigated sentences as smallest relevant units of linguistic communication but have begun to explore linguistic units beyond sentence boundaries and to analyze the structure of more complex units such as texts and discourses. These efforts have produced powerful discourse semantic frameworks, such as (S)DRT, Centering Theory, Accessibility Theory, and studies concerning the Question Under Discussion in texts and discourse (i.e. QUD or Quaestio) to name but a few (Walker 1998; Klein & Stutterheim 1987; Ariel 2001; Asher & Lascarides 2003; Stark et al. 2007; Onea 2016). Although there is still a lack of studies that apply these frameworks to (fictional) narrative texts, a number of recent studies within theoretical linguistics, historical linguistics, and linguistic typology deals with typical narrative phenomena and provides analysis of these phenomena in the frameworks mentioned above (see, for instance, the discussion of Free Indirect Discourse (FID) in the works of Schlenker 2004; Eckardt 2012, 2014; and Maier 2015 among others). Moreover, there is an increasing number of broad empirical and experimental studies, including corpus linguistic studies on linguistic change in discourse structures and narrative structures (see e.g. Petrova & Solf 2010 and Schlachter 2012) as well as psycholinguistic studies in the new field of experimental semantics and pragmatics investigating the processing and acquisition of (fictional) narrative texts (see e.g. Bortolussi & Dixon 2003; Burkhardt 2005; Carroll & Stutterheim 2011; Kaiser & Cohen 2012; and Kaiser 2015).

Since natural languages come in three different modalities, spoken languages, written languages and sign languages, linguistic studies on narration should not only focus on written language but do well to also integrate investigations on narration and discourse in spoken and sign language as well as on co-speech gestures, which typically accompany oral and signed discourse and narration. Especially new data from sign languages broaden the perspective and offer a more comprehensive

© 2018 John Benjamins Publishing Company

understanding of general linguistic foundations of narration across modalities. Another important aspect in this field is the fact that text and discourse structures in sign languages are increasingly investigated on a formally high theoretical level. Especially work on discourse referents and reference tracking as well as attitude and action role shift – two modality-specific strategies of presenting somebody's speech, thought and action in sign languages (for more detail, see Section 2 below) – has yielded interesting parallels with reference tracking and different kinds of reported speech such as FID and mixed quotation in spoken languages (see Quer 2005, 2011; Herrmann & Steinbach 2012; Lillo-Martin 2012; Schlenker 2016, 2017; Barberà 2015; Davidson 2015; Onea & Steinbach 2016 among others).

And finally, recent research on sign languages opens up an interesting new perspective on the interaction of manual and non-manual co-speech gestures with spoken and sign languages in narration and discourse. Gestures do not only fulfill important functions in structuring linguistic units in both modalities, they are also used to express the handling, location, movement, and size and shape of objects, topographic relations, speaker oriented (non-at-issue) meaning and specific aspects of reported speech and action (on the interaction of gesture and language see e.g. Kita & Özyürek 2003; Özyürek 2012; Müller et al. 2013; Loon et al. 2014; Goldin-Meadow & Brentari 2017).

The contributions in this collected volume investigate different aspects of narration in written, spoken and sign languages as well as multimodal narration including co-speech gestures in spoken and sign languages from various perspectives. In addition, they present and analyze various kinds of new empirical data from different languages and modalities and discuss many aspects of the acquisition, change, production, and processing of narrative structures. As will be shown below, each chapter makes an important contribution to a new interdisciplinary linguistic research area, which is still in its infancy. Although the chapters address quite different topics and can only add small new pieces of data and theoretical analysis to the general topic, we are convinced that the data and theories discuss in this collected volume provide a good basis for future research which hopefully gives us a more comprehensive picture of linguistic foundations of narration across modalities.

The following table gives an overview of the different modalities, linguistic aspects, and pragmatic topics addressed in the individual chapters.

As can be seen in Table 1, all eleven chapters of this volume can be grouped around three main topics, which count as top sellers in modern (linguistic) text analysis: information structure, anaphora resolution, and perspective. The first top seller, i.e. information structure, is addressed in detail in the first three contributions by Stutterheim and Carroll, Tomita, and Wilbur and Malaia. The following four chapters of Petrova, Speyer, Salem, Weskott, and Holler, as well as Barberà

Table 1. Topics that are discussed in the respective chapters of this volume

chapters	written languages	spoken languages	sign languages	gesture	language change	language acquisition	typology	theory	experimental linguistics	information structure	anaphora resolution	perspective
Wilbur and Malaia			+				+	+	+	+		
Tomita	+					+	+	+	+	+		+
Stutterheim and Carroll	+						+	+	+	+	+	
Petrova	+				+			+			+	
Speyer	+				+			+			+	
Salem, Weskott, and Holler	+							+	+		+	+
Zeman	+	+						+				+
Köder		+		+		+			+			+
Bressem, Ladewig, and Müller		+	+	+				+				+
Barberà and Quer			+	+			+	+			+	+
Herrmann and Pendzich			+	+			+					+

and Quer focus on different aspects of anaphora resolution. The last two chapters – Salem, Weskott and Holler, and Barberà and Quer – and the chapter of Tomita also address the third best seller, namely perspective, which is at the heart of the contributions of Zeman, Köder, Bressem, Ladewig and Müller as well as Herrmann and Pendzich.

The first two top sellers, information structure and anaphora resolution, are already well established in linguistic research (i.e. they are easily available at every linguistic bookshop). By contrast, perspective is not only one of the most recent (and least known) topics of linguistic analysis of discourse and narration but also the topic that is discussed in most of the chapters of this collected volume. Therefore, we focus on the third top seller in the remainder of this introduction and discuss different aspects of perspective shift in various kinds of narration in more detail in the following section. In the discussion, we are especially interested in FID and role shift, which are two prominent devices for shifting perspective. In Section 3, we then briefly summarize the main outcomes of the eleven contributions to this volume.

2. Perspective

The following passage in Example (1) taken from Jane Austen's *Emma* displays a typical example of FID (see also Zeman this volume). As a reader, one has the impression to have direct access to Emma's thoughts, whose story is told, that is, within FID, the point of view is shifted from the perspective of the narrator to the perspective of the protagonist.

(1)   *She could not be too soon alarmed, nor send for Perry too often. It was a pity, perhaps, that he had not come last night; for, though the child seemed well now – very well considering – it would probably have been better if Perry had seen it.*
[Jane Austen, *Emma*, example taken from Nikiforidou 2012: 180]

Such utterances are weird in many respects because they are *unspeakable* in the term of Banfield's (1982) seminal study *Unspeakable sentences. Narration and representation in the language of fiction*. This is not only because we deal with thoughts and not with utterances but also because of the contradictory combination of linguistic elements in FID, such as, for instance, the temporal deictic elements *last night* and *now* as opposed to the third person pronouns and the past tense. While *last night* and *now* (just as the speaker-oriented adverbials *perhaps* and *probably*) have to be interpreted with respect to the perspective of the character, the 3rd person pronouns and the past tense reflect the perspective of an external narrator. One might expect that such hybrid utterances, which combine two different contexts without any formal (grammatical or orthographic) marking, are ungrammatical or at least semantically anomalous. But neither the first nor the latter is the case. It seems to be perfectly fine to encode the perspectives of two speakers simultaneously within one single utterance.

The investigation of FID and related phenomena like focalization and interior monologue among others has a long tradition in literary studies where countless narratives concerning the specific manifestation of these phenomena and their implications for the interpretation of the narrative in question have been discussed.

However, beginning with Banfield (1982), FID has aroused the interest of formal linguistics. Since then, many researchers have tried to formulate a consistent theory of the semantic interpretation of FID accounting especially for its specific *mix* of indexical expressions. For the moment, two different lines of argumentation compete for the most adequate analysis of FID. On the one hand, there is the "context shift camp" with exponents like Schlenker (2004), Sharvit (2008), and Eckardt (2012, 2014). According to such analyses, the semantic interpretation of FID is relative to two different contexts, the context of the protagonist and the context of the narrator. Hence, FID triggers a shift in perspective from the narrator to the protagonist. To account for the behavior of indexical expressions in FID, context

shift approaches assume that languages have two different kinds of indexicals: personal pronouns and tense morphology are narrator-oriented and hence lexically specified to be linked to the context of the narrator. By contrast, temporal and local indexicals as well as speaker-oriented adverbials are protagonist-oriented, i.e. their reference is relative to the context of the protagonist. Differences in context-shift approaches result form the underlying type of discourse: While Sharvit assumes that FID is derived from indirect discourse, Schlenker and Eckardt consider FID to be a variant of direct discourse.

The rival camp is the (one-man) "camp unquotation" of Maier (2015, 2016), who argues that FID "is essentially a quotation of an utterance or thought, but with unquoted tenses and pronouns" (Maier 2015: 345). These approaches have not only been fruitfully applied to the analysis of spoken language narration, but also to role play in language acquisition (cf. Köder 2015) and to the analysis of attitude and action role shift in sign languages (for a more detailed discussion of the latter, see Schlenker 2016, 2017; Davidson 2015; Hübl 2016).

FID is one prominent means to trigger a context shift in literary narration. However, it is worth mentioning that questions of perspective shift are also highly relevant in everyday narration (Tannen 1989). Recent studies on the acquisition of reported speech as well as on modality-specific aspects of speech and thought report in sign languages show that perspective shift and perspective mix is not restricted to FID but can also be found in non-literary narration in both modalities. One prominent means of indicating perspective shift is so-called role shift, which is found in both spoken and sign languages. Role shift is a complex device in which the speaker or signer uses various linguistic as well as manual, non-manual and (in spoken languages) vocal gestural components to signal that he/she changed the point of view (Lillo-Martin 2012). Two aspects are important in this context: First, while role shift in spoken languages is basically gestural, action and attitude role shift in sign languages is integrated into the linguistic system and combines linguistic with gestural components (Herrmann & Steinbach 2012, see also Barberà & Quer, this volume, and Herrmann & Pendzich, this volume). The grammaticalization of gestural components in sign languages is possible because sign language and gesture use the same modality. As a consequence, sign languages have the unique property to integrate gestural elements into the linguistic system (Pfau & Steinbach 2011). Second, in sign languages, role shift can be used to express multiple perspectives in one utterance. Sign languages have the modality-specific property to realize various grammatical features simultaneously by using different articulators at the same time (Meier 2002). Similarly, different articulators can express different perspectives in role shift. In narratives, each hand can represent one protagonist, while at the same time, the face of the signer may represent a third protagonist and the whole body the narrator (see Herrmann & Pendzich, this volume). Hence, role shift

in sign languages is used to fulfill a similar function as FID in spoken languages – irrespective of the considerable formal differences between these two devices.

In sum, new studies on perspective shift provide evidence for the claim that the shiftability of indexical expressions is subject to modality-specific (and possibly also to language specific) constraints. In addition, perspective shift can be indicated by various linguistic and gestural means including speaker-oriented expressions, body shift, and speech imitation. And finally, at least in sign language narration, perspective mix can involve more than to two perspectives simultaneously. Hence, given these interesting properties of perspective shift across modalities and the exciting new developments in linguistic research, it comes as no surprise that most chapters of this volume address different aspects of FID and role shift in spoken and sign languages.

3.   Contributions to this volume

In the first contribution to this volume, **Ronnie Wilbur** and **Evie Malaia** offer a new method to experimentally study sentence prosody in short narratives in sign languages. So far, linguistic principles of prosody in signed texts have not been investigated with kinematic methods. Based on a discussion of various perceptual and production studies of fluency and prosody in American Sign Language (ASL), the authors present an experimental motion capture study that builds on prior kinematic studies. With this study, the authors pursue three main objectives. The first two objectives involve the confirmation of previous results. First, they aim at establishing the validity of the new method proposed in this chapter by confirming the results of the previous studies that have shown significant phrase final lengthening effects. The second objective comprises the validation of the function of increased velocity for stressed signs. And finally, the authors report new findings such as lengthening of signs in list conditions and prosodic effects that are related to the structure of narratives. The present study is thus an important step towards a systematic prosodic analysis of (different kinds of) narratives in sign languages. In addition, the method can fruitfully be applied to the analysis of co-speech gesture in spoken language narration thus providing new experimental data for multimodal communication across modalities (for motion capture in gesture research, see e.g. Schüller et al. 2017).

The prosodic organization of a text is just one task the signer or speaker is set with in narration or discourse. Text production also requires the signer or speaker to produce a coherent text. One important factor for coherence is the structural organization of information. Among other things, this involves the embedding and licensing of information at the micro structural level of the text by larger structures

established at the macro structural level. The organization of information at the micro and macro structural level is the topic of the two cross-linguistic studies presented in the second and third chapter. In her empirical study on narration, **Naoko Tomita** compares the use of two grammatical categories (aspect and point of view) in narrations of L1 and L2 speakers of two typologically different languages (Japanese and German). Although both groups use the same general principles to establish coherence in narration, differences in the grammatical realization of aspect and point of view in both languages also trigger differences in the organization of the narrative. Concerning aspect, Japanese speakers, unlike German speakers, make use of a grammaticalized verbal aspect to organize the temporal structure of narratives. By contrast, German speakers typically use lexical items such as adverbials or change-of-state verbs to reach the same goal. With respect to point of view, the author observes that German speakers, unlike Japanese speakers, do not take the protagonists point of view but tell the stories from their own perspective. This difference in perspective taking is accompanied by a difference in the way causal coherence is established. The study provides convincing evidence for the assumption that typological variation between languages has an impact on the way speakers organize narratives at the micro and macro structural level.

In the third chapter, **Christiane von Stutterheim** and **Mary Carroll** follow a similar line of argumentation. Their analysis is embedded in a Quaestio (or QUD based) framework, which assumes that the initial question underlying a text opens a space of alternatives, which have to be specified by the speaker or narrator in the text he/she produces. The linguistic implementation of these general macro level constraints is again subject to language-specific grammatical features at the lower micro structural level, that is, the availability of specific grammatical features in a language directly affects the linguistic micro structural organization of the corresponding text or narrative. In their empirical study, the authors compare two different genres, object descriptions and narratives, in two different languages, English and German. This design enables the authors to investigate different underlying questions at the macro structural level ('where is what?' in object descriptions and 'what happened to x at t_n?' in narratives) and the grammatical implementation in two different languages at the micro structural level. The evaluation of the data shows that a language with relatively fixed word order like English uses different strategies to answer the Quaestio as opposed to a language with relatively free word order like German. These differences concern especially the syntactic subject and the sentence-initial position. In German, spatial and temporal relations can easily occupy the sentence-initial position to ensure anaphoric linkage. By contrast, English frequently draws of the domains of entities given the restrictions on word order.

Two chapters of this volume focus on language change and explore the principles of anaphora resolution in Old High German (OHG). In the first chapter,

Svetlana Petrova investigates how the prominence of a discourse referent affects the selection of a certain kind of anaphora. The author presents two studies on the distribution of personal pronouns and different kinds of demonstrative pronouns. Both studies provide evidence for the assumption, that in OHG, discourse referents which are taken up more frequently in the following context are picked up by anaphora with a lower degree of lexical explicitness. By contrast, anaphora with a higher degree of explicitness are used to pick up discourse referents with a lower degree of persistence in the following context. Hence, the author shows that discourse prominence is an important factor constraining anaphora resolution in OHG. In addition, the prominence of a discourse referent is not only calculated on basis of the immediately preceding context but also on basis of the entire discourse, including the subsequent discourse (i.e. global prominence). This means that the choice of a particular anaphoric expression does not only depend on the preceding context but also affects the subsequent discourse.

In the second contribution that discusses historical data, **Augustin Speyer** investigates the distribution of personal and demonstrative pronouns in Otfrid's *Evangelienbuch*. His analysis of coherence relations and anaphora resolution in this text is embedded in the framework of centering theory. Although the OHG system is less sensitive to centerhood than Modern German (i.e. the choice of anaphoric expression does not directly depend on coherence relations such as continue, retain or shift), the analysis reveals certain tendencies for OHG. While in high coherent continue relations, backward looking centers are mainly realized as personal pronouns, in less coherent shift relations and in retain relations, they tend to be realized as demonstrative pronouns. The system of OHG thus resembles the one attested for Modern German. However, since the author also finds violations of this tendency, he concludes that, compared to Modern German, in OHG, centering is less important as a factor of anaphora resolution. In the history of German, the system operative in Modern German is attested at latest for Early New High German (ENHG). Both systems – ENHG and Modern German – share the following two properties: (i) Demonstrative pronouns are not associated with continue relations and (ii) in ambiguous reference, personal pronouns mark a higher ranked (preferred) coherent relation, while demonstrative pronouns mark a lower ranked coherent relation (Winter's rule).

FID has become a popular issue not only for theoretical linguists; psycholinguists have also started experimental investigations of how readers process FID. In previous experimental studies on FID (cf. Bortolussi & Dixon 2003), the methodological standards of modern experimental linguistics are, however, not met. Therefore, the two psycholinguistic experiments using both, on-line and off-line measures discussed by **Susanna Salem, Thomas Weskott** and **Anke Holler** present pioneering work in this field (cf. also Kaiser 2015 for an experimental study of

perspective shift in FID). Recall that FID is a linguistic means that induces readers to take over the protagonist's perspective. One possible prediction is that the referent of a protagonist exhibits an increased salience during reading if his/her thoughts are presented in FID, that is, reading times on an anaphoric expression referring to the protagonist should be lower when readers interpret the previous utterances as FID and hence shift into the perspective of the protagonist. In order to test this hypothesis, the authors use typical FID cues such as questions and discourse particles. Interestingly, the results do not provide evidence for the hypothesis that the narrative perspective influences the activation status of the corresponding discourse referent of the protagonist during reading. However, the FID cues have an influence on the perception of the narrative perspective.

Since FID seems to be restricted to narration and a defining property of certain kinds of narratives, **Sonja Zeman** uses the characteristics of FID and the different analysis of FID argued for in the literature to answer the important and very basic question: *What is a narration?* She points out that despite the increasing interest in the linguistic structure of narratives and the interpretation of these structures, the concept of narration itself has not yet been satisfactorily defined from a linguistic point of view. She concludes that the differentiation between narrator and protagonist in FID can be seen as a projection of the grammatical differentiation between speaker and observer, which is reflected on different linguistic levels in a recursive manner. She argues that the interaction of these layers is a basic property of narration and thus typical for narrative texts. Hence, each linguistic account of narration should at least include a thorough analysis of the grammatical means used to identify and track context and perspective shift.

As already mentioned above, FID is one prominent means to trigger a context shift but perspective shift is not restricted to FID. Speakers and signers may use various devices to indicate a shift of perspective in narration. Role shift is one prominent device that has received a great deal of attention in recent research. In this volume, four chapters deal with different aspects of role shift at the interface between gesture and language. Two chapters focus on spoken languages and two on sign languages. The first contribution in this field discusses role play in language acquisition. **Franziska Köder** compares the distribution of different perspectives in various kinds of speech report such as direct and indirect speech and role play. She argues that the first two kinds of reported speech (i.e. direct and indirect speech) differ from role play in three dimensions. First, in direct and indirect speech, the context shift is overtly indicated, i.e. various linguistic and metalinguistic devices are used to mark the utterance as speech report. By contrast, in role play, the external perspective of narration is not linguistically represented in the utterance. Second, the communicative intention of the speaker differs between role play on the one hand and direct and indirect speech on the other: pretending in role play versus

reporting in direct and indirect speech. And third, in role play, the speaker uses the gestural devices mentioned above to embody the protagonist, i.e. the speaker does not only report the utterance of the protagonist in a direct or indirect way but also gesturally demonstrates the linguistic and non-linguistic behavior of the protagonist (cf. also Davidson 2015 for a demonstration account of role shift in sign language). These three differences between role play on the one hand and (in)direct speech on the other can be related to the interpretative effects these different kinds of speech report trigger: While role play yields the strongest shift in perspective (i.e. identification with the protagonist), indirect speech only causes a weak shift. Direct speech can be located between role shift and indirect speech in this respect.

The second chapter that deals with multimodal narration in spoken languages is the one by **Jan Bressem, Silva Ladewig,** and **Cornelia Müller**. The authors investigate gestural aspects of perspective shift in spoken language narration. The study provides evidence that speakers have available different kinds of a (gestural) depiction of the action described, i.e. a speaker can give a semantically reduced multimodal description of the event described in the narration or they use various gestural articulators to give a semantically more complex multimodal description of the event. Hence, multimodal utterances vary on a continuum of semiotic complexity depending on the bodily involvement of gestural articulators. Consequently, the authors argue for an expansion of the notion of 'character viewpoint gestures' to the notion of 'multimodal action depiction from a character viewpoint'. The study shows that speakers use in principle the same multimodal devises signers use in action role shift to depict actions. One difference, which we already mentioned above, is that sign languages, unlike spoken languages, integrate these devises into the linguistic system. Therefore, we expect quantitatively and qualitatively more bodily involvement of gestural articulators in sign language narration.

The last two chapters of this collected volume turn to perspective shift and anaphora resolution in sign languages narration. Although both studies use the same kind of naturalistic data, i.e. a set of signed narratives, and investigate similar constructions such as role shift and classifiers, they focus on different aspects. **Gemma Barberà** and **Josep Quer** investigate the function of classifiers for reference tracking in narratives. They argue that in sign language narration, role shift cannot only be used to reproduce the words of a protagonist but also the actions. The first kind of role shift, which is used for reported speech, is called attitude role shift. The second one, which is a grammatical-gestural device to report the actions of protagonists, is called action role shift (or constructed action). The authors argue that these two kinds of role shift share properties with both direct and indirect quotation in spoken languages. Nonmanual features such as body shift and change of facial expression mark a context shift and indicate who is signing or acting. As already discussed in the contribution by Salem, Weskott, and Holler, a shift of perspective

may affect reference tracking and anaphora resolution. In this chapter, Barberà and Quer focus on the dynamics of discourse referents in role shift in Catalan Sign Language (LSC). A qualitative analysis of classifier constructions in LSC's versions of the Aesop's fables shows that classifiers, just like pronouns, are used to establish anaphoric chains in narration. The authors argue for a dynamic semantic analysis of coherence, which also includes semantic relations among referential expressions. Classifiers are treated as associative anaphoric expressions in role shift. Thus, role shift has two functions in this context: First, it licenses the coreference between a classifier and a previously introduced discourse referent. Second, it makes the corresponding discourse referent the most prominent one at this stage of discourse. Classifier constructions in role shift provide thus new evidence for a dynamic semantic approach of accessibility in discourse.

The final contribution by **Annika Herrmann** and **Nina-Kristin Pendzich** deals with perspectivation and gestures in action role shift in German Sign Language (DGS) and combines questions addressed in the previous three sections on gesture and sign language. Based on the same set of Aesop's fables, the authors show that a signer can express multiple perspectives under role shift. Expressing multiple perspectives simultaneously seems to be a modality-specific property of sign languages. It is well-known that sign languages can express various grammatical features simultaneous because they can use different manual and nonmanual articulators independently (Meier 2002). In action role shift, a signer can use these articulators to represent different discourse referents and the corresponding perspectives simultaneously. These articulators interact in a complex way to represent multiple shifts of perspectives in a discourse segment. As already discussed in the previous chapter by Barberà and Quer, classifier constructions are one prominent means of representing multiple perspectives. In addition, facial expressions can be used to give additional non-at-issue information about the attitude of one of the protagonists. Interestingly, both classifier constructions and facial expressions incorporate gestural components, that is, sign languages systematically make use of gestural components in action role shift to indicate multiple perspectives. As already mentioned above, sign languages and gestures use the same modality. Therefore, sign languages can more easily interact with gestures than spoken languages. The analysis of the DGS fables reveals that role shift can be classified at a continuum from pure grammatical quotation (i.e. attitude role shift) to highly gestural action demonstration (i.e. action role shift). Consequently, the authors argue for a new analysis of these two kinds of role shift to account for the modality-specific complex interaction of grammar and gesture and for the possibility to express multiple perspectives in sign language narration.

References

Asher, N. & Lascarides, A. 2003. *Logics of Discourse*. Cambridge: CUP.

Ariel, M. 2001. Accessibility theory: An overview. In *Text Representation: Linguistic and Psycholinguistic Aspects* [Human Cognitive Processing 8], T. Sanders, J. Schilperoord & W. Spooren (eds), 29–87. Amsterdam: John Benjamins.

Banfield, A. 1982. *Unspeakable Sentences. Narration and Representation in the Language of Fiction*. Boston MA: Routledge & Kegan Paul.

Barberà, G. 2015. *The Meaning of Space in Sign Language. Reference, Specificity and Structure in Catalan Sign Language Discourse* [Sign Languages and Deaf Communities 4]. Berlin: De Gruyter Mouton and Ishara Press.

Bortolussi, M. & Dixon, P. 2003. *Psychonarratology. Foundations for the Empirical Study of Literary Response*. Cambridge: CUP.

Burkhardt, P. [Schuhmacher]. 2005. *The Syntax-discourse Interface: Representing and Interpreting Dependency* [Linguistik Akuell/Linguistics Today 80]. Amsterdam: John Benjamins.

Carroll, M. & von Stutterheim, C. 2011. Event representation, time event relations, and clause structure: A crosslinguistic study of English and German. In *Event Representation*, E. Pederson & J. Bohnemeyer (eds), 68–83. Cambridge: CUP.

Davidson, D. 2015. Quotation, demonstration, and iconicity. *Linguistics & Philosophy* 38(6): 477–520.

Eckardt, R. 2012. Particles as speaker indexicals in free indirect discourse. In *Particles*, L. Hogeweg & E. McCready (eds), 99–119. Special issue of Sprache und Datenverarbeitung.

Eckardt, R. 2014. *The Semantics of Free Indirect Discourse. How Texts Allow us to Read Minds and Eavesdrop*. Leiden: Brill.

Goldin-Meadow, S. & Brentari, D. 2017. Gesture, sign and language: The coming of age of sign language and gesture studies. *Brain and Behavioral Sciences* 39: 1–17.

Herrmann, A. & Steinbach, M. 2012. Quotation in sign languages. A visible context shift. In *Quotatives. Cross-linguistic and Cross-disciplinary Perspectives* [Converging Evidence in Language and Communication Research 15], I. Buchstaller & I. van Alphen (eds), 203–228. Amsterdam: John Benjamins.

Hübl, A. 2013. Role shift, indexicals and beyond – New evidence from German Sign Language. In *Proceedings from the 13th Meeting of the Texas Linguistics Society*, L. C. Geer (ed.), 1–11. Austin TX: University of Texas.

Hübl, A. 2014. Context shift (im)possible – Indexicals in German Sign Language. In *Proceedings of ConSOLE XXI*, M. Kohlberger, B. Kate & E. Dutton (eds), 171–183. Leiden: Sole.

Hübl, A. 2016. *Pointing in Context. Semantic and Pragmatic Aspects of Indexicals in Sign Language Role Shift*. PhD dissertation, University of Göttingen.

Kaiser, E. 2015. Perspective-shifting and free indirect discourse: Experimental investigations. In *Proceedings of Semantics and Linguistic Theory 25 (SALT 25)*, S. D'Antonio, M. Moroney & C. Rose Little (eds), 346–372. https://journals.linguisticsociety.org/proceedings/index.php/SALT/issue/view/132

Kaiser, E. & Cohen, A. 2012. *Free Indirect Discourse and perspective-taking*. Poster presented at the annual conference Architectures and Mechanisms for Language Processing (AMLaP), Riva del Garda, Italy. [https://dl.dropboxusercontent.com/u/7489928/amlap2012.pdf]

Kita, S. & Özyürek, A. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48: 16–32.

Klein, W. & von Stutterheim, C. 1987. Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte* 109: 163–183.

Köder, F. 2015. *Between Direct and Indirect Speech. The Acquisition of Pronouns in Reported Speech*. PhD dissertation, University of Groningen.

Lillo-Martin, D. 2012. Utterance reports and constructed action. In *Sign Language*, R. Pfau, M. Steinbach & B. Woll (eds), 365–387. Berlin: Mouton de Gruyter.

Loon, van E., Pfau, R. & Steinbach, M. 2014. The grammaticalization of gestures in sign languages. In *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill & J. Bressem (eds), 2133–2149. Berlin: De Gruyter Mouton.

Maier, E. 2014. Mixed quotation: The grammar of apparently transparent opacity. *Semantics and Pragmatics* 7(7): 1–67.

Maier, E. 2015. Quotation and unquotation in free indirect discourse. *Mind & Language* 30(3): 345–373.

Maier, E. 2016. *A plea against monsters*. Ms, University of Groningen.

Meier, R. P. 2002. Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In *Modality and Structure in Signed and Spoken Languages*, R. P. Meier, K. Cormier & D. Quinto-Pozos (eds), 1–25. Cambridge: CUP.

Müller, C., Ladewig, S. H. & Bressem, J. 2013. Gestures and speech from a linguistic perspective: A new field and its history. In *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill & J. Bressem (eds), 55–81. Berlin: De Gruyter Mouton.

Nikiforidou, K. 2012. The constructional underpinnings of viewpoint blends. The Past + now in language and literature. In *Viewpoint in Language. A Multimodal Perspective*, B. Dancygier & E. Sweetser (eds), 177–197. Cambridge: CUP.

Özyürek, A. 2012. Gesture. In *Sign Language*, R. Pfau, M. Steinbach & B. Woll (eds), 626–646. Berlin: Mouton de Gruyter.

Petrova, S. & Solf, M. 2010. Pronominale Wiederaufnahme im ältesten Deutsch. In *Historische Textgrammatik und historische Syntax des Deutschen*, A. Ziegler & C. Braun (eds), 339–363. Berlin: De Gruyter.

Pfau, R. & Steinbach, M. 2011. Grammaticalization in sign languages. In *Handbook of Grammaticalization*, B. Heine & H. Narrog (eds), 681–693. Oxford: OUP.

Quer, J. 2005. Context shift and indexical variables in sign languages. In *Proceedings of Semantics and Linguistics Theory 15 (SALT 15)*, E. Georgala & J. Howell (eds), 152–168. Ithaca NY: CLC Publications.

Quer, J. 2011. Reporting and quoting in signed discourse. In *Understanding Quotation*, E. Brendel, J. Meibauer & M. Steinbach (eds), 277–302. Berlin: Mouton de Gruyter.

Schlachter, E. 2012. *Syntax und Informationsstruktur im Althochdeutschen. Untersuchungen am Beispiel der Isidorgruppe*. Heidelberg: Winter.

Schlenker, P. 2004. Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind & Language* 19(3): 279–304.

Schlenker, P. 2016. Super monsters I. Attitude and action role shift in sign language. *To appear in Semantics and Pragmatics*.

Schlenker, P. 2017. Super monsters II. Role shift, iconicity and quotation in sign language. *To appear in Semantics and Pragmatics*.

Schüller, D., Beecks, C., Hassani, M., Hinnell, J., Brenger, B., Seidl, T., & Mittelberg, I. 2017. Automated pattern analysis in gesture research: Similarity measuring in 3D motion capture models of communicative action. *Digital Humanities Quaterly* 11(2).

Sharvit, Y. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31(3): 353–395.

Stark, E., Leiss, E. & Abraham, W. (eds) 2007. *Nominal Determination. Typology, Context Constraints, and Historical Emergence* [Studies in Language Companion Series 89]. Amsterdam: John Benjamins.

Steinbach, M. & Onea, E. 2016. A DRT analysis of discourse referents and anaphora resolution in sign language. *Journal of Semantics* 33: 409–448.

Tannen, D. 1989. *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: CUP.

Walker, M., Joshi, A. & Prince, E. 1998. Centering in naturally occurring discourse: An overview. In *Centering Theory in Discourse*, M. Walker, A. Joshi & E. Prince (eds), 1–28. Oxford: Clarendon Press.

A new technique for analyzing
narrative prosodic effects in sign languages
using motion capture technology

Ronnie B. Wilbur and Evie Malaia
Purdue University

The present paper addresses prosody at a sentence level analysis within short narratives, developing a novel method based on a combination of data. Our first objective, then, is to establish the validity of our new method by confirming the findings of previous reports on Phrase Final Lengthening. Our second objective is to further extend the validity of the new method by confirming a prior claim regarding the role of increased velocity for signs that are stressed in sentences. Finally, we report new results in the form of patterns across signs within sentences using our new method.

1. Introduction

This research program addresses a simple question: What is fluent signing? The answer to this question is, however, not at all simple to find. One problem that interferes with investigating the answer is the absence of technology and established analytical methodology comparable to what is available for speech (cf. PRAAT). Motion capture equipment and recording techniques are still in development, and there is no standard analytical procedure or software. The procedures that we will describe in this paper are unlike those used by anyone, including us, up to this point. Furthermore, these procedures are still in development, and we have taken the extra precaution of demonstrating that the results we obtain match those that have been reported by earlier methods. There are two main differences between present and prior methods: (1) prior methods have targeted measurements at the individual sign level even when the productions are obtained from carrier phrases or three-sign sentences, whereas present methods aim toward capturing the prosody of signs in more naturalistic sentences (sentence-effects) and sentences in narratives (narrative-effects); and (2) prior methods required multiple (usually five)

productions of each stimuli by each signer in order to obtain measures of variability, whereas the present method obtains large numbers of short fluent scripted narratives (48 for the present report) produced only once by a single signer. The goal, then, is not to determine absolute values for any of the measurement variables but rather to identify the relationships between the variables and the prosodic behaviors of interest, from which a model can be built.[1] Subsequent testing of the model would be the target of future work.

1.1   Perceptual studies of prosody and fluency

Two types of perceptual studies relevant to prosody and fluency have been conducted so far. In one, judges are shown videos of signers and are asked either (1) to rate the fluency of signers on a variety of criteria or (2) to determine prosodic boundaries based on the cues produced by signers.

   With respect to signer fluency, Kantor (1978) began the study of sign fluency by asking the question of how well fluent signers could be separated from non-fluent signers, using the assumption that Deaf native signers were, by definition, fluent signers (assuming no obvious motoric disturbances). In contrast, the status of hearing native signers and both Deaf and hearing L2 (second language, post-puberty) signers was not known. At the same time, it was not known which groups of signers would be able to separate fluent from non-fluent signers, nor how much an effect there would be on ratings from the signers' hearing status. The two groups that were most easily identified were the native Deaf signers and the L2 hearing signers. In contrast, Deaf L2 signers were correctly identified as deaf but rarely as L2. Kantor interpreted these results as indicators of Deaf community status ("in" vs. "out"). Among the criteria mentioned by judges were facial expression, use of mime, exaggerated mouthing (early oral training), and rhythm of signing, including speed, fluidity and use of space. The L2 deaf and hearing judges performed considerably below the native groups, suggesting that they were less attuned to the salient characteristics.

   Lupton (1998) followed up to try to determine more specifically what criteria contribute to fluent signing. She observed that raters valued how the signs were produced - smooth and steady instead of choppy, hesitant, and jerky. Other motoric aspects, such as rate of signing or movement amplitude, did not correlate with fluency ratings. She also reported that less fluent signers were more likely to use excessive mouth movements, and to show less eye contact, facial expression, and body movements. These latter aspects – eye contact, facial expression and body

---

1. Techniques for normalizing sign data to permit absolute value comparison across signers are under development (Grosvald 2009; Russell et al. 2011).

movements – are correlated with syntactic abilities as well, and thus it is important to note that the concept of fluency is not simply a rhythmic/motoric notion.

Attempts to capture sign language prosody have been for the most part based on video data (Brentari 1998; Boyes Braem 1999; Brentari & Crossley 2002; Weast 2008; Wilbur 1994, 1997, 2000, 2010, 2011a,b; Wilbur & Patschke 1998), and those that involve experimental tasks. Descriptively, of most relevance is the fact that the Prosodic Hierarchy holds for ASL (Brentari 1998; Sandler & Lillo-Martin 2006). However, this topic is outside of the focus of the current paper; summaries can be found in Tang et al. (2010), Sandler (2012), and Ormel & Crasborn (2012).[2] Experimental techniques to capture prosody have included tapping (Allen, Wilbur & Schick 1991), signing at different rates (Grosjean & Lane 1977; Wilbur 2009), and cue judgments (González 2011; Brentari et al. 2011; Brentari, Nadolske & Wolford 2012).

González (2011) presented ASL signing to four groups of viewers: ASL signers, Hong Kong Sign Language (HKSL) signers, non-signers, and second language (L2) learners of ASL. Viewers were asked to identify prosodic boundaries. Differences were found across all four groups in terms of both accuracy and reaction times. Non-signers and L2 signers were more accurate when a broader range of boundary cues were available, and HKSL signers were more accurate than ASL signers, indicating that knowing the language can result in distraction due to language processing.

Brentari, Nadolske, and Wolford (2012) also identified the relative strengths of various cues with respect to the perception of boundaries by native and L2 ASL signers and non-signers: sign duration, presence of holds, transition duration between signs, pause duration (hold plus transition), blinks, drop hands, brow position change, head position change, and torso position change. Of these, only sign duration has so far been investigated kinematically.

1.2   Production studies of fluency and prosody

With respect to fluency of production, it is often assumed that adult learners of a sign language already have the motor coordination needed to fluently produce signed sentences and that the primary challenge to fluid performance is getting the right signs in the right sentence position in the appropriate structure for the

---

2.   It should be noted however that in Sandler (2012), as well as in Nespor & Sandler (1999) and Sandler & Lillo-Martin (2006), the position is taken that intonation is carried by the upper face while the hands produce the text. In contrast, Wilbur (2009) argued that the presence of upper face articulators is driven by semantics (showing the restriction of dyadic operators or the scope of monadic operators) and that these facial articulations are, once present, subject to the same prosodic processes that affect the hands (rhythmic phrasing, lengthening, stress).

context. However, Lupton and Zelaznik (1990) demonstrated that this is not the case. Normal young adult learners of ASL did not achieve appropriate bilateral coordination of their left and right hands in two-handed signs until about 12 weeks into their first semester course.

Another common misconception is that speed of signing is critical to perceived fluency. It is true that there is good correspondence between syntactic and prosodic breaks (Grosjean & Lane 1977), and that pause duration and other prosodic markers depend on signing rate (Wilbur 2009), but ratings of fluency are not dependent on signing rate (Lupton 1998).

1.3   Prior kinematic work

1.3.1   *Prior motion capture work showing sign lowering*
Previous motion capture work focuses on analysis at the single sign level, with the target sign placed inside of a carrier phrase in order to ensure that the target sign can be properly extracted from the motion capture stream. This procedure enabled investigation of lowering of the target sign's place of articulation resulting from the height effects of the previous or following sign, as well as from signing rate (Mauk & Tyrone 2008, 2012; Tyrone & Mauk 2010). Other research has contributed to understanding the possible nature of motor disorders in sign language (Tyrone 2007; Tyrone, Atkinson, Marshall & Woll 2009).

Tyrone et al. (2010) provides a task-dynamic analysis of sign production, as did Wilbur (1990) and McDonald et al. (2016), which aims at determining the relative temporal components of signs. Along similar lines, Ormel and Crasborn (2012), using CyberGloves, compare the transitions between signs with the lexical movements of signs and report lower velocity in the transitional movements. They suggest that this cue helps the viewer distinguish lexical from transitional movements of the hand (recall that the hand is always visible, whether making a sign or not).

1.3.2   *Prior motion capture work showing verb class differences*
Our own previous research at the single sign level varied from the standard procedures somewhat in each study (cf. Malaia et al. 2012ab). One study focused on verbs produced by four deaf ASL fluent signers, but unlike standard tasks, they did not produce each stimulus multiple times in the same condition (Malaia & Wilbur 2012). The study design derived analytical power from larger numbers of diverse stimuli (40 different signs) recorded once per signer in each of four conditions across four signers. This allowed us to report the kinematic differences across well-represented linguistically meaningful classes of signs rather than at the level of individual signs. Thus, signers produced each target sign once in four different conditions (isolation, carrier phrase, carrier sentence medial, and carrier sentence

final), and there were 40 different target verb signs chosen from two semantically distinct groups. As observed by Wilbur (2003, 2008, 2010), event structure differences in the meaning of the verbs, that is, whether the verb denotes a telic event (one with an end-state, such as HIT) or an atelic event (such as TRAVEL), are reflected in the formation of the verb signs, with telics having sharper end-marking than atelics.[3] The kinematic results show that the end-points of telic signs are marked by significantly greater deceleration than atelics and that this difference is not affected by prosodic processes related to sentence position (Phrase Final Lengthening).[4] Thus, the analysis compared the two groups of diverse signs against each other across different conditions.

To determine whether the recruitment of kinematic features occurs in other sign languages for the same purposes, we conducted another study with one native hearing signer of Croatian Sign Language (*Hrvatskom Znakovnom Jeziku*, HZJ) who produced translated stimuli in the same four conditions in separate sessions (Malaia, Wilbur & Milković 2013). Thus, the HZJ data represents one signer, albeit on five different days and over 240 productions per day. These studies are driven by the Event Visibility Hypothesis (Wilbur 2003, 2008), which proposes that such use of kinematic features should be universal to sign languages by the grammaticalization of physics and geometry for linguistic purposes.

HZJ differs from ASL in how telic and atelic verbs signs are related. In ASL, they are simply different unrelated roots, although there are some signs which can have their movement modified to allow them to alternate between denoting telic and atelic events. In HZJ, there are also those types of signs, but HZJ has a systematic process that ASL lacks: although all verbs with telic meaning have the marking of deceleration, ASL has no regular morphological process to produce an alternation between two forms of a verb from one stem. In contrast, looking at 200 HZJ aspectual pairs of verbs (400 total verbs), Milković (2011) found that

---

3. This end state marking could be thought of as being like a suffix that combines with the verb in the same way that the past tense shows up on English verb 'walk', which although written 'walked' is pronounced /walkt/, that is, as a single syllable. The addition of the end state suffix in telic verb forms does not make a second syllable, but simply joins the existing syllable at its end.

4. As a reminder of the physics: velocity (v) = distance (d) divided by time/duration (t). Acceleration is the change in velocity over time. Acceleration can be discussed in two ways: (1) technically, deceleration is negative acceleration, that is, change in velocity in the negative direction (meaning slowing down, not 'changing direction of movement'); as such, *deceleration* is not used as a term; and (2) more generally, acceleration can be used to mean 'speeding up' and deceleration to mean 'slowing down'. We will use acceleration in labeling our variable MinA, the minimum acceleration will be a negative number, and 'greater minA' will more negative. For ease of exposition, we will use deceleration as a general term when not specifically referring to the variable itself. The reader is also reminded that all of these variables are completely visible in signing.

they formed three groups. The largest systematically altered the properties of sign movement: the same root would appear with shorter, sharper movement for telic as compared to atelic. The second group did not allow alternation of telic and atelic signs from the same roots but instead used phrasal sequences of several types: verb plus a separate aspectual sign, quantification of the internal argument, or use of verbal complements. These are comparable to expressions used in English, for example, atelic 'to run' in 'Mary ran' can be telic in 'Mary ran to the store'. The third group displayed pairs using suppletive stems parallel to ASL. Our motion capture study focused on the alternating group, although data for the suppletive group have already been collected.

Like ASL, HZJ also showed significant differences in deceleration between verb classes. This observation supports the EVH claim that the end-point in event structure is kinematically manifested as 'end-marking' in sign production, whether such marking is unique to each sign root (i.e. lexical), as in ASL, or used productively throughout the verbal paradigm, as in HZJ. One clear difference between the two languages is peak velocity. For HZJ, the peak velocity was greater in telic signs as compared to atelic ones, and the effect of position was not significant (Figure 1). In contrast, in ASL peak velocity is affected both by verb type *and* phrase position. One possible interpretation of this difference between the two sign languages is that grammaticalization of event structure in HZJ makes the parameter of peak velocity robust to prosodic effects. A related possibility is that in ASL, peak velocity is used to indicate stress (Wilbur 1999), whereas currently there is no information concerning the marking of stress in HZJ, which may use a different motion variable or other type of marking (face/head/body). Further research is required to fully understand this kinematic difference between the two languages. This is the first cross-linguistic motion capture confirmation that specific kinematic properties of articulator motion are grammaticalised in other sign languages to express linguistic features.
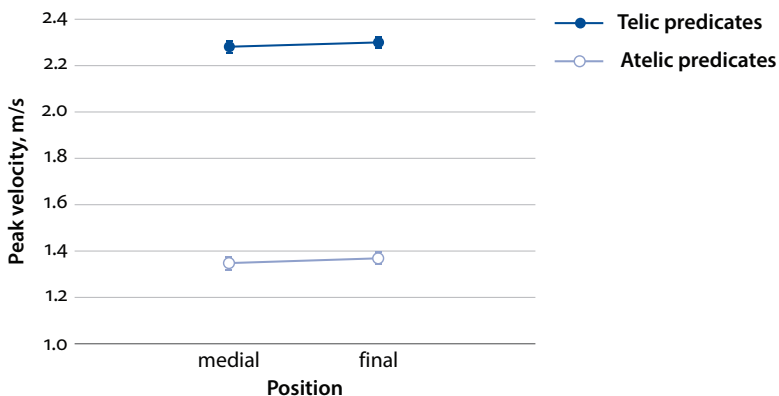
Figure 1. HZJ peak velocity showing no effect of phrase final lengthening

1.3.3   *Prior motion capture work showing phrase final lengthening*

It should also be noted that both of these prior studies establish sentence level kinematics insofar as the phrase position of the target signs compares final with non-final effects. In ASL, the variables of duration and peak velocity are both affected by Phrase Final Lengthening; duration is expected from the very name of the process. The fact that peak velocity is also affected makes it an unreliable candidate to mark telicity, leaving deceleration as the main marker for ASL. HZJ also shows Phrase Final Lengthening effects on duration; however peak velocity is unaffected, whereas deceleration is affected (Figure 2). This marks a sharp contrast between the two languages.
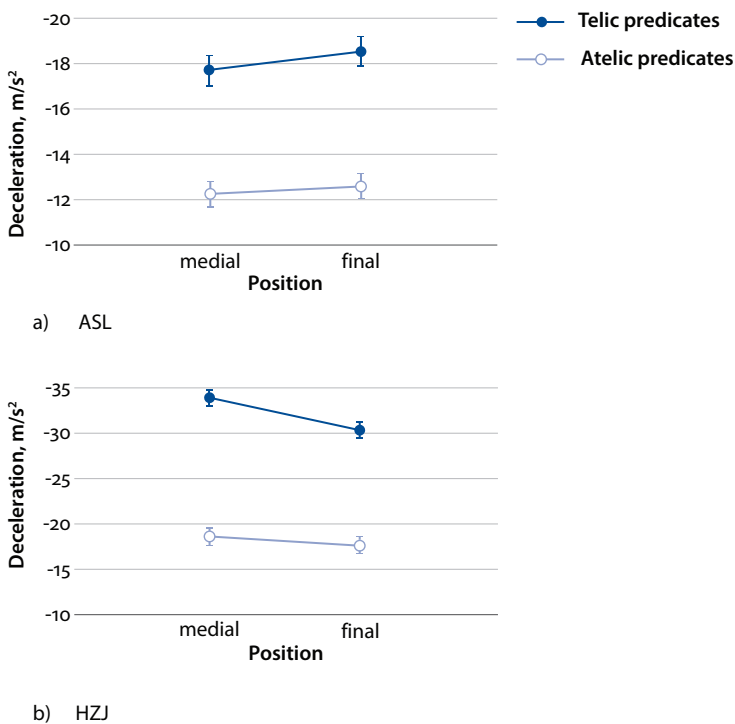
Figure 2. Deceleration values for telic and atelic signs in medial and final position (note differences in maximum value of deceleration on Y-axis)

In Malaia, Wilbur, and Milković (2013), we introduced another calculated measure, the Ratio of the telic-atelic slopes. This gives an overall indication of how fast the movement is coming to a stop in the two types of signs. To be clear, MinA is the value of the greatest deceleration achieved during the production of a sign's movement, that is, the fastest rate of change in velocity while slowing. The slope is a more general measure, as it is calculated as the average slowing from MaxV

to MinV (the change in velocity divided by how long it takes to get from MaxV to MinV). In both languages, the slope for telics is larger than for atelics. However, a comparison of the Ratios of telic slope to atelic slope reveals that the HZJ telic slopes are close to twice those of atelics (final telic to atelic: −.016/ −.009 = 1.78; medial telic to atelic: −.019/ −.010 = 1.9), whereas for ASL the difference is much smaller (final telic to atelic: −.011/ −.007 = 1.57; medial telic to atelic: −.010/ −.008 = 1.25). This shows that HZJ also uses deceleration as a marker, but that deceleration is secondary to peak velocity in this language.

1.3.4   *Prior motion capture work showing stress marking*

The final prosodic marking that has been investigated kinematically using single target sign methods is stress. Wilbur (1999) summarises results of a motion capture study of 13 ASL signers producing target signs in carrier phrases for four relevant contexts: target sign is (1) stressed and final, (2) stressed and medial, (3) unstressed and final, and (4) unstressed and medial. Given that ASL prefers prominence in final position (Wilbur 1999), these carrier phrases were necessarily artificial albeit acceptable ASL, requiring unsigned context to make them semantically coherent. For example, signers were willing to produce "YOU MUST READ BOOK" once it was agreed that typically one would add another MUST at the end but that it was possible without it, albeit "Englishy," or ELLEN MUST READ BOOK, [not Susan]. Other examples took advantage of word order possibilities in ASL: YOU READ BOOK MUST versus YOU MUST READ BOOK. The results provided the earliest instrumental (non-video) documentation of significant Phrase Final Lengthening effects on duration. In addition, whether target signs were stressed or unstressed did not affect duration; only peak velocity was significantly affected by stress. This means that signers do not have to make a four-way distinction in duration, that is, a different average duration for each combination of stress and phrase position. Instead, position is shown by duration, and stress is shown by velocity. Furthermore, displacement (how far the hands travel) is not affected by either of these variables, meaning that it is available to vary as needed to permit the distinctions in duration and velocity (see again fn. 2).

2. The present study

The present paper addresses prosody at a sentence level analysis within short narratives, developing a novel method based on a combination of data. Our first objective, then, is to establish the validity of our new method by confirming the findings of previous reports on Phrase Final Lengthening. Our second objective is to further extend the validity of the new method by confirming a prior claim regarding the role of increased velocity for signs that are stressed in sentences. Finally, we report new results in the form of patterns across signs within sentences using our new method.

To understand the structure of the current study, it is necessary to review an earlier stress study that did not involve motion capture. Wilbur and Schick (1987) scripted 48 different narratives (3 sentences each) that put 24 target signs in stressed and unstressed positions. For example, the sign DIE is targeted as stressed in (1a) and unstressed in (1b).

(1) a. SHOCK ME. DISCOVER GOOD FRIEND **DIE**. THINK HEART-ATTACK. NOT-KNOW … SEEM SICK HE, NOT-KNOW ME.
   "I was shocked to discover that my good friend died. I think it was a heart attack but I'm not sure. Apparently he was sick but I didn't know."

   b. MY FRIEND MARRY AGAIN. WIFE FIRST **DIE** LONG-AGO. NOW HAVE WIFE.
   "My friend married again. His first wife died a long time ago. Now he has a wife again."

Fourteen Deaf ASL signers produced the narratives, which were recorded. To confirm that the stress target signs were indeed stressed, two judges separately watched all 14 signers producing all narratives. Each judge circled the corresponding gloss on a judgment sheet for any sign that appeared to them to be stressed. The present study uses the same 48 narratives and the stress judgments but in novel ways to be described below.

Before proceeding, we should report the results observed by Wilbur and Schick. When stressed signs were compared to their unstressed counterparts, the stressed signs were seen to be set off from the surrounding unstressed signs by 'sharper transition boundaries', to be higher in the signing space, and produced with apparent increased muscle tension. These differences in production were determined by skilled ASL linguists watching 30 fps videotape repeatedly. Subsequent instrumental studies have attempted to better understand what produces these visual effects.

Turning now to the present study, we report one set of data recorded during an entire afternoon of motion capture recording of one Deaf ASL signer. During that time, none of the stimuli were repeated unless the signer felt that the previous production was not right. Our goal is to develop techniques and to demonstrate

their validity and reliability that will permit us to analyse large quantities of data, such as collected that afternoon and on other occasions, representing fluent ASL in scripted as well as unscripted narratives and stories. Ideally we can characterise signing that is fluent as well as signing that is not fluent (for example, the few examples that were rejected and repeated during the recording sessions). To be able to deal with the small number of non-fluent examples, we need techniques that do not require multiple repetitions to obtain.

For the present study, we analyse the kinematic data from one ASL signer in a motion capture suit who produced the same 48 narratives used for Wilbur and Schick (1987). We combine that data with a stress WEIGHT variable derived from the older study, as well as with additional linguistic variables describing the signs and sentences in which they occurred.

2.1   Procedure

2.1.1   *Motion capture method for current report*

A Deaf fluent ASL signer was recorded producing the same 48 short narratives used in Wilbur and Schick (1987). The signer wore a Gypsy 3.0 motion capture suit, and the data about XYZ positions of all markers were collected at the rate of 60 fps by 6 specialty cameras mounted in a circle on the ceiling. With these cameras, no markers are ever occluded from recording, eliminating the recurrent problem of missing data in standard motion recording setups. A simultaneous video recording at 30 fps rate was made with an NTSC video camera on a tripod outside the motion capture recording field. The positional data from the marker on the right wrist, tracking the movement of the dominant signing hand, was used for the analysis. These procedures parallel those used in Malaia and Wilbur (2012a) and Malaia, Wilbur, and Milković (2013).

2.1.2   *Coding for non-kinematic linguistic and perceptual variables*

The video was imported into ELAN annotation software and aligned using audio markers and the T-pose (the signer standing with hands extended to the sides at shoulder level) at the beginning and end of each recording section. The video was then annotated in ELAN by a series of signers in the Purdue ASL Lab. The first round of annotation added the gloss of each sign and marked the beginning and end of each sign following procedures established by Green (1984), assuming the first frame of recognition of the sign-initial handshape as the beginning of each sign, and either the point of contact, or maximal distance traveled by the hand, as the end of the sign. Thus, the onset and the ending of each sign were defined linguistically based solely on the video cues, without access to kinematic variables.

The remaining five rounds of annotations not only checked accuracy of the first round but also added additional information. Each sign was coded for the following variables: narrative number (Story), sentence number within each narrative (Sentence), gloss number within each sentence (Gloss), whether the sign was originally target to be stressed or not (Stress), and the position of each sign within its sentence/phrase (Position). For Position, each sign was coded (1or 0) for each possibility: Initial, Final, neither (Medial), or both (List). Our inclusion of List (as both Initial and Final in its phrase) is a novel contribution of this study; signs so coded can be either in an actual list or e.g. a single sign Topic. In (2) TAP-SHOULDER is phrase final, ICE-CREAM is both phrase initial and phrase final, as is POP.

(2) ALWAYS POSS-1 DAUGHTER TAP-SHOULDER, ICE-CREAM, POP. IX-1 BLEW-UP, TELL-IX-3 LATER, DON'T BOTHER IX-1.
"My daughter is always bothering me for ice cream, soda pop. I told her "Later! Don't bother me.""

In addition to these linguistic measures, we calculated a perceptual measure of stress called Weight through the following procedure. We started with the (archived) original stress judgments of the two judges from the Wilbur and Schick (1987) study; these judges had separately watched 14 signers producing the same 48 narratives. The judges had each circled the corresponding glosses printed on paper if the sign appeared to them to be stressed. We calculated a "stress score" by totaling the number of times each sign was judged to be stressed (2 judges × 14 signers) and converted the total for each sign to percent to yield the variable Weight.

2.1.3   Coding for kinematic variables

The beginning and end time points for each sign were exported from ELAN and processed in MATLAB to extract speed and acceleration profiles for each sign from the recorded kinematic files. The following kinematic metrics were identified or calculated for each sign following the procedure determined by Malaia and Wilbur (2012a):

a. duration in milliseconds (DURation);
b. peak instantaneous speed achieved (MaxV);
c. the local minimum speed following the peak speed (MinV);
d. the percent of sign movement elapsed to the moment where peak speed occurred (% elapsed dur), which is also the point at which deceleration starts;
e. subsequent minimum deceleration (MinA)
f. the slope from MaxV to MinV (Slope)

An example of each variable for the ASL sign ʜɪᴛ is given in Figure 3 (from Malaia & Wilbur 2012a); the variables are labeled on the displacement, speed, and acceleration profiles for the sign.
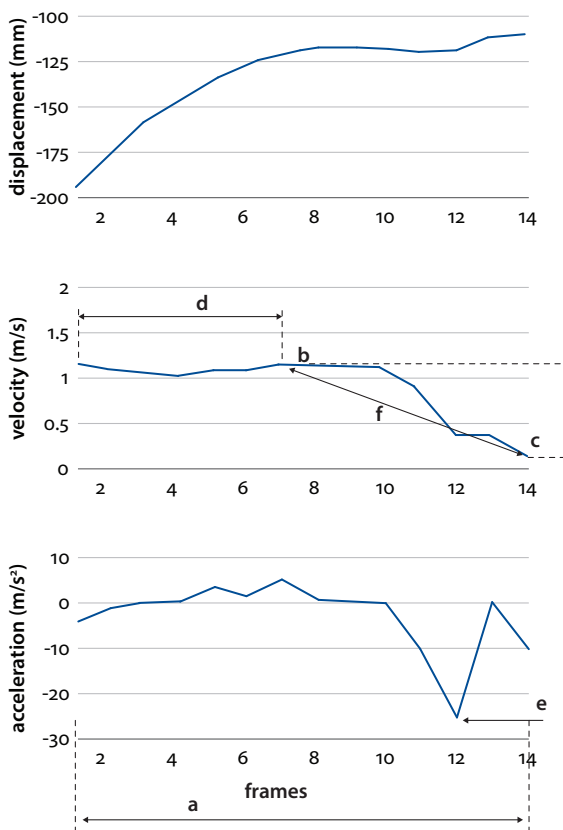
Figure 3. Kinematic variables for ASL ʜɪᴛ. The variables are labeled on the displacement, velocity, and acceleration vectors: (a) sign duration; (b) maximal velocity (MaxV); (c) minimum velocity following peak velocity (MinV); (d) percent of sign elapsed to peak velocity (% elapsed); (e) minimal acceleration (MinA) following peak velocity; (f) overall slope of deceleration from peak velocity to the following minimum velocity (slope).

2.2   Analyses

Variables were combined from multiple sources: ELAN exported output, the calculated variable Weight, and kinematic measures derived from Matlab analyses of motion capture data. Several statistical analyses were performed using SPSS 20. These

included paired t-test for comparing the stressed and unstressed sign pairs and multiple analysis of variance (MANOVA) with regression and post-hoc analyses.

Our analysis began with 694 signs in 144 sentences across the 48 paragraphs, or approximately 14.5 signs per narrative. Signs with maximal speed occurring on the last or next to last frame were discarded from analysis (21% of cases); these were the cases where contact occurred at the end of the sign, but both hands kept moving together briefly after contact; this situation resulted from using Green's (1984) definition for determining the sign end, which perhaps cuts the end short.[5] This resulted in 548 signs being available for analysis.

2.3   Results

2.3.1   *Kinematic analysis by sentence position*

The analysis by sentence position focused first on validating the procedure by determining whether the data provided evidence for Phrase Final Lengthening. Towards this end, the duration of signs (DUR) was analysed against position (Initial, Medial, Final, and List); means and standard deviations are given in Table 1 (Figure 4). Position had an overall effect on Duration ($F(3,544) = 39.629$, $p < .001$). The mean duration of Final signs is 69.6% longer than Initial signs and 52% longer than Medial signs.[6] These results indicate that there is Phrase Final Lengthening in our data, across 144 sentences and 48 narratives. This finding provides a measure of validity for this method.

Table 1. Sign duration by sentence position

Position	Mean (ms)	Std. deviation	N
Initial	461.53	280.22	157
Medial	513.53	267.76	178
Final	782.45	440.32	167
List	913.58	431.31	46
Total	614.17	381.97	548

---

5.   This exclusion was enforced by a filter in SPSS to prevent miscalculation of the slopes (avoiding division by 0 or 1). If each variable were analysed separately, these cases could have been omitted just for the slope analysis, but it was deemed preferable to begin by using a stable number of cases across all variables.

6.   Bear in mind that Initial signs here does not include Topic signs, which are included in the List position to be discussed.
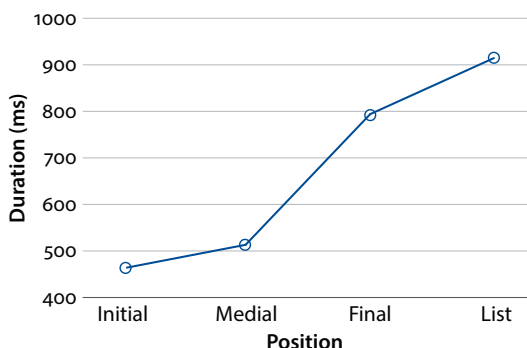
Figure 4. Sign duration by sentence position

In addition to Phrase Final Lengthening, there is also an interesting lengthening of signs in List condition. As indicated earlier, signs coded as List are those which are both Initial and Final in their phrases.[7] There were 48 signs in this group, which includes 14 that are also the first sign in their sentences, indicating that they are Topics.[8] These signs are 16.8% longer than those that are coded only as being in Final position. This extra length in List position is part of what accounts for their prosodic distinctiveness as reflected in the fact that their Weight score (how stressed they were perceived to be by the judges watching these narratives produced by 14 ASL signers not wearing motion capture suits) was, like Final position, significantly different from Initial and Medial positions (Bonferroni-adjusted post-hoc testing).

   Knowing what signs in List situations look like kinematically compared to those in regular sentence positions may lead the way to developing a procedure for characterizing non-fluent signing. Presumably, someone with Parkinson's disease or other motor disfluencies might show list-like productions in greater number and in places inside sentences that would not be appropriate given their actual sentence position. This procedure needs to be developed further, but the goal is to be able to describe the difference between the prosody of natural sentences and simple lists of signs in a row. If we can capture this difference, and if it is based primarily on duration, we might be able to develop a measure that does not require motion capture equipment to calculate.

---

7. List was not an option for coders however; they were only coding Initial and Final positions. Both List and Medial were subsequently calculated variables: List if both Initial and Final were equal to one, and Medial if both Initial and Final were zero.

8. That is, they are sentence initial and also final in their own phrase. There are no sentences that consist of only one sign.

Our analysis by sentence position also reveals sentence position effects on some of the remaining variables. Both % elapsed to maxV and MinA were significantly affected (F(3,544) = 2.718, $p < .05$ for %elapsed to MaxV; 4.527, $p < .005$ for MinA). MaxV approached significance (F(3,544) = 2.387, p = .068). There were no other main effects for sentence position. Figures 5–9 display the individual variables mapped against sentence positions. The graphs illustrate that Final differs from Initial and Medial, and that sometimes List groups with Final and sometimes it is distinct. Our next planned study includes multiple lists of signs (in addition to natural narratives), so we will return to this problem in the future to get a better picture of the kinematics of list behavior.
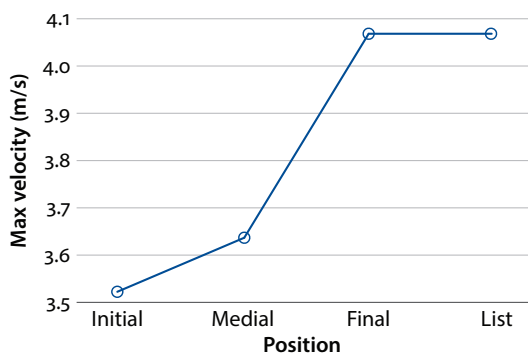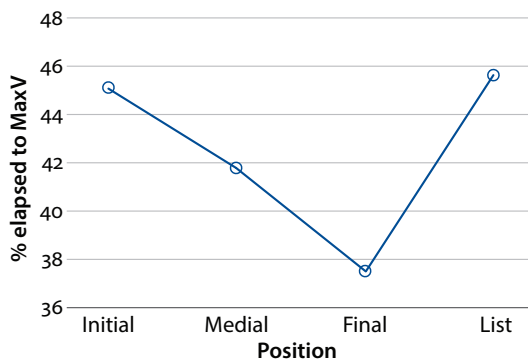

Figure 5. MaxV by sentence position

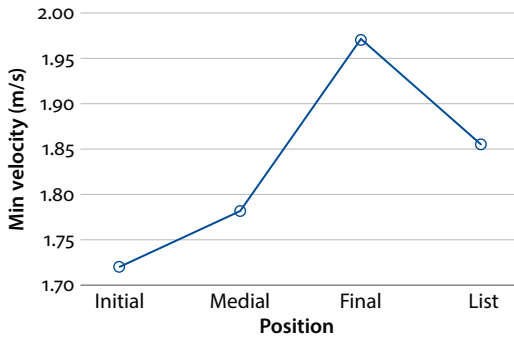**Figure 6.** Percent elapsed sign movement to MaxV by position

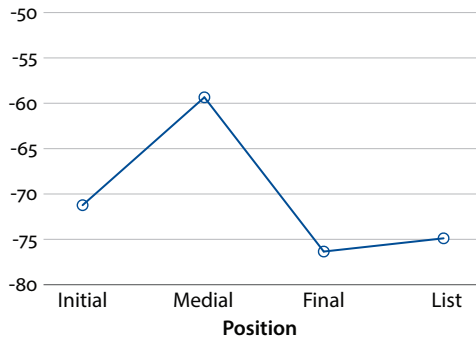**Figure 7.** Minimum velocity after MaxV

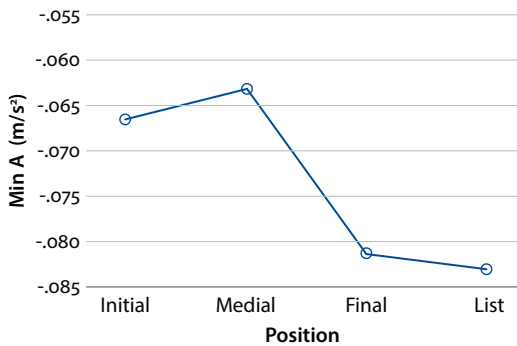**Figure 8.** Slope from MaxV to MinV by position

Figure 9. MinA by position

2.3.2 *Kinematic analysis of stress*

A particular advantage of the current data set is that the 48 narratives represent 24 pairs of signs which are targeted as stressed in one of the narratives and as unstressed in the other narrative. Given that the original scripting of the narratives was done long before motion capture was readily available, there was no concern in the scripting for trying to make the two narratives in the pair as similar as possible except for the target's stress status. Given what is now known about stress and focus in ASL (Wilbur 1997, 1999), such scripting could not be accomplished beyond simple albeit manipulated carrier phrases. Thus we are analyzing data in which both the stressed and unstressed target signs are surrounded by different lexical items in different syntactic structures and possibly even in different sentence sequences in the narratives. In other words, we are looking at signing that is close to natural spontaneous signing which, even though scripted, was nonetheless scripted by a native ASL signer. Ideally, techniques developed to analyse this data set will eventually be generalizable to natural signing (at least as natural as can be achieved with motion capture sensors on and/or video cameras recording).

Because we have pairs of signs, we are able to perform paired t-tests, which are the strongest tests for differences.[9] Each sign is tested against itself for each variable: Duration, MaxV, % Elapsed Duration, MinV, slope, and MinA (Table 2). We provide the data on Weight to show that indeed stressed signs were judged as stressed more often (giving a higher weight score) than unstressed signs.

The only variable that differs significantly between stressed signs and their unstressed counterparts is MaxV, the peak velocity achieved during the sign's movement.[10] This finding is consistent with the prior motion capture result reported by Wilbur (1999) and provides another measure of validity for our current method. Having said that, it is clear that unless the data is set up to permit this type of comparison, paired t-tests will not be of particular use in capturing generalizations about prosodic structure over narratives.

Having looked at the stronger test, we proceed to investigate the contribution of each variable to stress marking using multiple variable analysis of variance

---

9. If we wanted to find out if people generally have different size feet, we measure everyone's left foot and right foot and determine if there is a difference between them (pair each left with each right foot), then analyse across individuals to get an average difference in size, rather than comparing the mean of everyone's left feet with the mean of everyone's right feet because this latter calculation is subject to greater variance (suppose there are many more taller people than shorter people, and perhaps taller people have a greater size difference than shorter people).

10. Note that MinA came close to significance. As a kinematic variable, MinA was not studied prior to Malaia and Wilbur (2102), so this is its first appearance in a study related to stress.

Table 2. Paired t-test for stressed-unstressed targets[*]

	Paired differences			t	df	Sig.
	Mean	**SD**	**SE**			
Duration	82.457	559.462	122.085	0.675	20	0.507
MaxV	1.909	2.862	0.625	3.056	20	0.006
% elapsed to Max V	−13.206	52.912	11.546	−1.144	20	0.266
MinV	0.947	3.619	0.789	1.199	20	0.245
Slope	−0.527	1.413	0.392	−1.346	12	0.203
MinA	−0.031	0.070	0.015	−1.985	20	0.061
Weight	−0.296	0.245	0.054	5.526	20	0.000

* Note the reduced number of degrees of freedom. Some of the unstressed targets were so unstressed that they were not produced by the signer at all, which eliminates that pair from processing. In addition, the restriction concerning where MaxV is found with respect to the end of the sign (not less than 2 frames from the end), eliminated several pairs from being analysed for that variable.

(MANOVA) conducted through the generalised linear method (GLM).[11] The variables are tested against the two groups of signs: target stress signs or unstressed signs. This analysis revealed significant effects of Duration and MinA on stress, whereas MaxV only approached significance.[12] One problem with conducting analysis this way, however, is that we already know that position in the sentence significantly affects duration, and also that many of the signs that occur in final position were not targeted as stressed in the original narratives, but they are included in the unstressed group even if they were produced as stressed. A second problem is that there were originally only 24 signs that were targeted to be stressed, so the unstressed signs greatly outnumber the target stressed ones (524 to 24), creating a large imbalance in the numbers in each group. Two additional analyses could be proposed: (1) to include both Stress and Position in the analysis, which addresses the first problem, but not the second; and (2) to look at a different analysis altogether in which the Weight variable (judgments of stress productions) is used rather than the targeted (intended) stressed signs (Stress). This method addresses both problems because all signs have a Weight score.[13]

---

11. This is the analysis in which everyone's left foot is analysed against everyone's right foot rather than in pairs by individual.

12. Slope also approaches significance.

13. There are however many signs which have a Weight score of zero. This results from either (a) the judges never circled it when watching the original 14 signers producing it, or (b) the sign was not included in the original scripted narratives and therefore was not judged on the answer sheet if it was present or it was not produced then but was produced by the signer wearing the motion capture suit.

First we show that Weight addresses both the position contribution and the small number of stress targets problems. This discussion will also allow us to demonstrate the behavior of Weight, a perceptual measure, as a useful variable in motion capture analysis.

We first conduct an analysis of variance of Weight against Stress and Position. As predicted, main effects are significant for both Stress ($F_{(1,540)} = 63.320$, $p < .001$) and Position ($F_{(3,540)} = 3.103$, $p < .03$), as is their interaction ($F_{(3,540)} = 2.645$, $p < .05$). First, signs which were stress targets have a mean weight score of .708 (max = 1.00; SE .065), whereas those that were not have a mean weight score of .183 (SE .01). Thus, Weight clearly distinguishes those signs that were targeted for stress as a separate group from those that were not (recall that Weight was not determined for motion capture or this signer but from previous signers and from videorecordings). Second, different sentence positions showed different weight scores: Initial .490 (SE .101), Medial .457 (SE .059), Final .337 (SE .03), and List .496 (SE .053); post-hoc testing shows that the only significant difference is List compared to the other positions. These numbers at first appear surprising because they seem to decrease as we go from initial to medial to final in contrast to what we might expect given our knowledge of stress behavior in ASL (Wilbur 1999). That is why it is necessary to look at the interaction between Stress and Position. These data are given in Table 3.

Table 3. Weight scores by stress by position

Position	Unstressed	Stressed
Initial	.124 SE .016	.857 SE .202
Medial	.128 SE .015	.786 SE .117
Final	.157 SE .016	.518 SE .058
List	.323 SE .031	.670 SE .101

The question we then ask is what kinematic variables contribute to the perception of stress as reflected by the Weight score. For this we use a regression of Weight against the other kinematic variables. A linear regression analysis shows what we expect – Weight scores are most affected by whether the target sign is stressed or not ($t = 8.967$, $p < .001$), with position effects contributed by sign duration ($t = 5.424$, $p < .001$). When duration is removed using regression analysis (filtering out position effects), Weight is significantly correlated with MaxV, slope, and MinA.[14] The analysis also reveals that MaxV is significantly correlated with all the other variables except duration. Because of these correlations, MaxV does not appear as

---

14. This provides a possible explanation for the appearance of MinA as a nearly significant variable in the paired t-test results.

a significant predictor of Weight when duration is included in the mix. As one can
see, unraveling the relationships among the kinematic variables requires careful
planning, as well as multiple post-hoc analyses. Clearly establishing the relation-
ships would not have been possible if we had not already conducted the paired t-test
and the earlier controlled carrier phrase study.

2.3.3   *More general test of relationships among measured variables and weight*
Now we turn to what can be characterised beyond the single sign and single sentence
level. As mentioned above, one goal is to be able to characterise the difference be-
tween natural signed sentences and simple lists of signs. Another goal is to be able to
characterise narrative as contrasted with a simple list of sentences. As a step in that
direction, we begin by looking for effects that are related to whether the sign occurs
in the initial, medial or final sentence of each narrative (SentCount), in addition to
preserving the information about the position where it occurs in its own sentence.

 We will report two illustrations that indicate the possibility that the medial
sentence may be an environment where other processes may be occurring, or al-
ternatively, where other processes are not occurring. From these, we can develop
testable hypotheses: (1) In a narrative, initial sentences have different prosody from
those that follow; and (2) In a narrative, the final sentence has different prosody
from those that precede. Testing of these hypotheses requires narrative data with
more than one medial sentence, thus, longer narratives than the current data set
provides. One thing we see is a difference in Final and List duration in the middle
sentence that does not occur in the first or last sentence (Figure 10). In the first and
last sentence, the durations of Final and List signs are relatively close to each other,
whereas in the second sentence, List appears to be much longer. Thus, we can say
that there is no obvious effect of sentence sequence on sign duration overall, but
there is a possible middle sentence effect on Final and List signs.
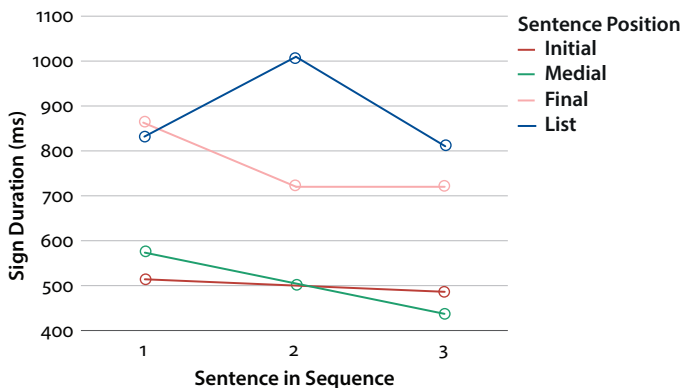
Figure 10. Sign duration by sentence position by sentence position in narrative

The second illustration of possible sentence sequence effects in narratives comes from analysis of MaxV (peak velocity), which as we have noted is a primary carrier of stress. We can rephrase the question as one of whether there is a difference in how stress is marked as a function of where the sentence is located in a narrative. This analysis looks at MaxV as a function of Position by Sentence in Sequence shown by whether the sign was targeted for stress or not. This analysis, displayed graphically in Figure 11, again shows that List is more variable than other positions. It also shows a larger variance in MaxV when the signs are stress targets. The MaxV patterns are also somewhat different for first, middle, and last sentences in sequence in the narratives. Perhaps the biggest point highlighted by this analysis is that ana-lyzing natural signing (or at least naturalistic signing from native signer scripted stimuli) is not the same thing as analyzing carefully controlled constructed stimuli. Note that in the stress target column, there are no signs in initial position in the first sentence that were targeted for stress, that is, one does not start a narrative with the
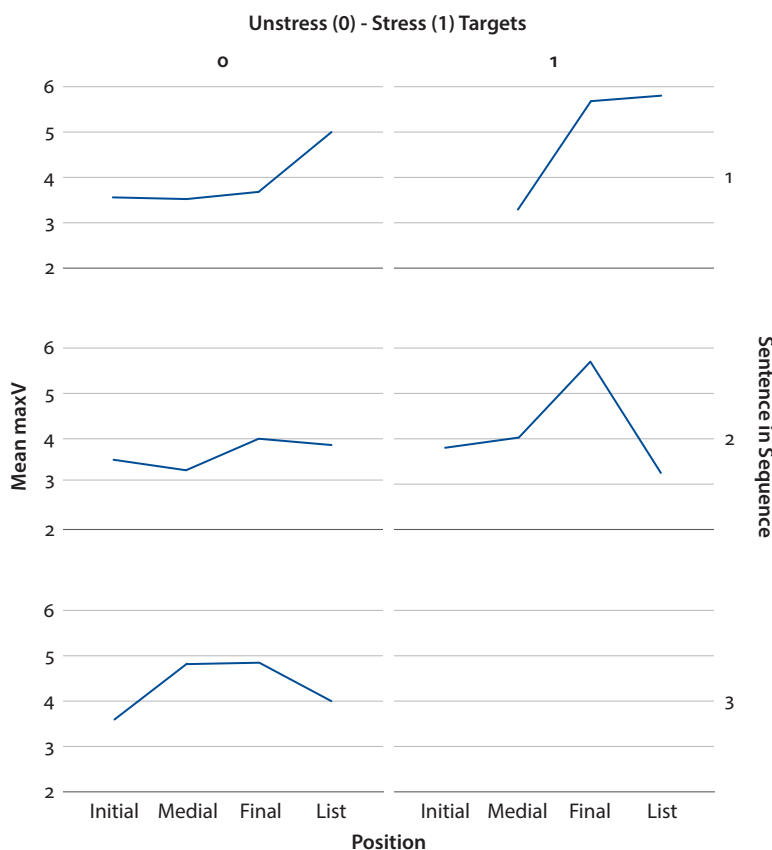
Figure 11. MaxV by position by sentence in sequence by stress

absolute first sign as stressed. Similarly, there were no signs in the last sentence of any of the narratives that were targets for stress, resulting in a blank graph. This tells us that when narratives are scripted to feel naturalistic (the original goal in Wilbur and Schick 1987), there are certain places where stressed signs do not feel natural: the beginning and the end of a narrative. In the case of the beginning, it is acceptable to have a sign be stressed in the initial sentence but not in the initial position of that sentence. In the case of the end, it appears that the entire last sentence may be off limits. These observations are suggestive, not definitive, and require different data and methods for subsequent analysis.

2.4   Summary

This study has used motion capture data to reconfirm previous analysis of linguistic stress as marked by increased peak velocity in ASL. The only variable that differs significantly between stressed signs and their unstressed counterparts is the peak velocity achieved during the sign's movement. This finding then provided a measure of validity for development of our newer method, which relies only a large amount of data collected from a single signer with each signed utterance signed only once, rather than the traditional five times. Our newer analysis revealed a significant contribution of sign duration and minimum peak acceleration (deceleration) to linguistic stress, with peak velocity only approaching significance. Once the additional contribution of phrase final position (duration lengthening) was taken into consideration by using the Weight variable, a perceptual measure, we were able to show that Weight addresses both the position contribution and the small number of stress targets problems. Finally, we have taken first steps towards understanding the prosodic effects of signed utterances in narratives and contributed the notion of List position as a potentially separate prosodic category.

3.   Conclusion

Prior kinematic work has established sentence-level effects such as Phrase Final Lengthening and stress marking, albeit in carrier phrase or single sentence stimuli. Larger narratives have rarely been addressed in production and not with kinematic analysis afforded by motion capture capabilities. The analyses provided in this paper have focused on demonstrating the potential of using different techniques for approaching motion capture data for linguistic purposes. Traditional techniques focus on analyzing variations in production across multiple repetitions per stimulus item across multiple signers. We have shown, with the analysis of duration and

Phrase Final Lengthening, that single productions of sufficient multiple stimuli by a single signer can be used to obtain comparable power and the same results as the traditional methods. We have also shown, using paired-t tests of target signs in stressed and unstressed contexts, that we can arrive at the same results as carrier phrases using traditional methods. In addition, our approach permits analysis of relatively more natural signing in longer narratives. From this, we are able to start the process of looking at the effects of sentence position within narratives as a possible way of capturing notions of fluent signing beyond single signs in carrier phrases or single sentences. An additional contribution of this approach has been the identification of "list" items, those which are the only sign in their phrase. Further characterization of this type of sign may enable us to develop measures of sentence prosody compared to 'signs in a string', and will also help to capture the difference between fluent signing and learner and movement-disrupted signing. Finally, the direction we are headed eventually will enable us to describe characteristics of different genres of signing – stories, instructions, explanations, spatial layouts, constructed actions, art-sign – as well as non-signing gesture.[15] We have previously suggested that sign language research can contribute techniques and models to gesture research (Wilbur & Malaia 2008; Malaia 2014, 2017), and have started to see more interaction in this domain (Pfeiffer 2013; Pfeiffer et al. 2013; Kroger et al. 2011; Malaia et al. 2016, 2017).

Acknowledgements

---

15. However, we must end with some caveats. Our current analyses are across all narratives but ignore narrative structure effects that may occur in our scripted narratives – this is not the analysis we want to end up with. We also want our analysis to capture 'within sentence within narrative', that is 48 separate analyses of three sentences in a row (what we have now is the average of 48 sentences across narratives, not the same thing). That cannot be done by current methods, so new approaches must be identified. But we do have a snapshot of within-sentence and within-narrative position effects.

References

Allen, G. D., Wilbur, R. B. & Schick, B. 1991. Aspects of rhythm in American Sign Language. *Sign Language Studies* 72: 297–320.

Boyes Braem, P. 1999. Rhythmic temporal patterns in the signing of deaf early and later learners of Swiss German Sign Language. *Language & Speech* 42: 177–208.

Brentari, D. 1998. *A Prosodic Model of Sign Language Phonology*. Cambridge MA: The MIT Press.

Brentari, D. & Crossley, L. 2002. Prosody on the hands and face: Evidence from American Sign Language. *Sign Language & Linguistics* 5: 105–130.

Brentari, D., Nadolske, M. & Wolford, G. 2012. Can experience with co-speech gesture influence the prosody of a sign language? Sign language prosodic cues in bimodal bilinguals. *Bilingualism: Language and Cognition* 15(2): 402–412.

Brentari, D., González, C., Seidl, A. & Wilbur, R. B. 2011. Sensitivity to visual prosodic cues in signers and nonsigners. *Language & Speech* 54: 49–72.

González, C. 2011. *Perception of Prosody in American Sign Language*. PhD dissertation, Purdue University.

Green, K. 1984. Sign boundaries in American Sign Language. *Sign Language Studies* 42: 65–91.

Grosjean, F. & Lane, H. 1977. Pauses and syntax in American Sign Language. *Cognition* 5: 101–117.

Grosvald, M. A. 2009. *Long-distance Coarticulation: A Production and Perception Study of English and American Sign Language*. PhD dissertation, University of California at Davis.

Kantor, R. 1978. Identifying native and second-language signers. In Sign Language Research, R. B. Wilbur (ed.), *Special issue of Communication and Cognition* 11(1): 39–55.

Kröger, B. J., Birkholz, P., Kannampuzha, J., Kaufmann, E. & Mittelberg, I. 2011. Movements and holds in fluent sentence production of American Sign Language: The action-based approach. *Cognitive Computation* 3(3): 449–465.

Liddell, S. K. 1977. *An Investigation into the Syntactic Structure of American Sign Language*. PhD dissertation, University of California, San Diego.

Lupton, L. K. 1998. Fluency in American Sign Language. *Journal of Deaf Studies and Deaf Education* 3(4): 320–328.

Lupton, L. K. & Zelaznik, H. 1990. Motor learning in sign language students. *Sign Language Studies* 67: 153–174.

Malaia, E. 2014. It still isn't over: Event boundaries in language and perception. *Language and Linguistics Compass* 8(3): 89–98.

Malaia, E. 2017. Current and future methodologies for quantitative analysis of information transfer in sign language and gesture data. *Behavioral and Brain Sciences* 40.

Malaia, E., Borneman, J. D. & Wilbur, R. B. 2016. Assessment of information content in visual signal: analysis of optical flow fractal complexity. *Visual Cognition* 24(3): 246–251.

Malaia, E., Borneman, J. D. & Wilbur, R. B. 2017. Information transfer capacity of articulators in American Sign Language. *Language and Speech*. doi: 10.1177/0023830917708461.

Malaia, E., Ranaweera, R., Wilbur R. B. & Talavage, T. M. 2012. Neural representation of event structure in American Sign Language: fMRI comparison of cortical activations in deaf signers and hearing non-signers. *Neuroimage* 59: 4094–4101.

Malaia, E. & Wilbur, R. B. 2012a. Kinematic signatures of telic and atelic events in ASL predicates. *Language & Speech* 55(3): 407–421.

Malaia, E. & Wilbur, R. B. 2012b. Telicity expression in the visual modality. In *Telicity, Change, and State: A Cross-categorial View of Event Structure*, V. Demonte & L. McNally (eds), 122–136. Oxford: OUP.

Malaia, E., Wilbur, R. B. & Milković, M. 2013. Kinematic parameters of signed verbs at the morpho-phonology interface. *Journal of Speech, Language and Hearing Research* 56: 1677–1688.

Malaia, E., Wilbur, R. B. & Weber-Fox, C. 2013. Event end-point primes the undergoer argument: Neurobiological bases of event structure processing. In *Studies in the Composition and Decomposition of Event Predicates*, B. Arsenijević, B. Gehrke & R. Marin (eds), 231–248. Dordrecht: Springer.

Mauk, C. E. 2003. *Undershoot in Two Modalities: Evidence from Fast Speech and Fast Signing*. PhD dissertation, University of Texas at Austin.

Mauk, C. E. & Tyrone, M. E. 2008. Sign lowering as phonetic reduction in American Sign Language. In *Proceedings of the 2008 International Seminar on Speech Production*, R. Sock, S. Fuchs & Y. Laprie (eds), 185–188. <http://issp2008.loria.fr/proceedings.html>

Mauk, C. E. & Tyrone, M. E. 2012. Location in ASL: Insights from phonetic variation. *Sign Language & Linguistics* 15(1): 128–146.

McDonald, J., Wolfe, R., Wilbur, R. B., Moncrief, R., Malaia, E., Fujimoto, S., Baowidan, S. & Stec, J. 2016. A new tool to facilitate prosodic analysis of motion capture data and a data-driven technique for the improvement of avatar motion. In *Proceedings of the Seventh Workshop on the Representation and Processing of Sign Languages: Corpus Mining (LREC-2016)*, 153–158, Portorož, Slovenia.

Milković, M. 2011. *Verb Classes in Croatian Sign Language (HZJ): Syntactic and Semantic Properties*. PhD dissertation, University of Zagreb, Croatia.

Nespor, M. & Sandler, W. 1999. Prosody in Israeli Sign Language. *Language and Speech* 42: 143–176.

Ormel, E. & Crasborn, O. 2012. Prosodic correlates of sentences in signed languages: A literature review and suggestions for new types of studies. *Sign Language Studies* 12(2): 279–315.

Pfeiffer, T. 2013. Documentation of gestures with motion capture. In *Body – Language – Communication*, C. Müller, A. Cienki, E. Fricke, H. S. Ladewig, D. McNeill & S. Tessendorf (eds), 857–868. Berlin: De Gruyter Mouton.

Pfeiffer, T., Hofmann, F., Hahn, F., Rieser, H. & Röpke, I. 2013. Gesture semantics reconstruction based on motion capturing and complex event processing: A circular shape example. In Proceedings of the SIGDIAL 2013 Conference, Maxine Eskenazi, Michael Strube, Barbara Di Eugenio & Jason D. Williams (eds), 270–279. Metz: ACL. <https://aclanthology.info/volumes/proceedings-of-the-sigdial-2013-conference>

Russell, K., Wilkinson, E. & Janzen, T. 2011. ASL sign lowering as undershoot: A corpus study. *Laboratory Phonology* 2(2): 403–422.

Sandler, W. 2012. Visual prosody. In *Sign language*, R. Pfau, M. Steinbach & B. Woll (eds), 55–76. Berlin: De Gruyter Mouton.

Sandler, W. & Lillo-Martin, D. 2006. *Sign Language and Linguistic Universals*. Cambridge: CUP.

Tang, G., Brentari, D., González, C., & Sze, F. 2010. Crosslinguistic variation in prosodic cues. In *Sign languages*, D. Brentari (ed.), 519–542. Cambridge: CUP.

Tyrone, M. E. 2007. Simultaneity in atypical signers – Implications for the structure of signed language. In *Simultaneity in Signed Language: Form and Function* [Current Issues in Linguistic Theory 281], M. Vermeerbergen, L. Leeson & O. Crasborn (eds), 317–335. Amsterdam: John Benjamins.

Tyrone, M. E., Nam, H., Saltzman, E., Mathur, G. & Goldstein, L. 2010. Prosody and movement in American Sign Language: A task-dynamics approach. In *Speech Prosody 2010 Conference Proceedings*. <http://speechprosody2010.illinois.edu/papers/100957.pdf>

Tyrone, M. E., Atkinson, J. R., Marshall, J. & Woll, B. 2009. The effects of cerebellar ataxia on sign language production: A case study. *Neurocase* 15(5): 419–426.

Tyrone, M. E. & Mauk, C. E. 2010. Sign lowering and phonetic reduction in American Sign Language. *Journal of Phonetics* 38(2): 317–328.

Weast, T. 2008. *Questions in American Sign Language: A Quantitative Analysis of Raised and Lowered Eyebrows*. PhD dissertation, University of Texas, Arlington, U.S.A.

Wilbur, R. B. 1990. An experimental investigation of stressed sign production. *International Journal of Sign Linguistics* 1(1): 41–59.

Wilbur, R. B. 1994. Eyeblinks and ASL phrase structure. *Sign Language Studies* 84: 221–240.

Wilbur, R. B. 1997. A prosodic/pragmatic explanation for word order variation in ASL with typological implications. In *Lexical and Syntactic Constructions and the Construction of Meaning*, Vol. 1 [Current Issues in Linguistic Theory 150], M. Verspoor, K. D. Lee & E. Sweetser (eds) 89–104. Amsterdam: John Benjamins.

Wilbur, R. B. 1999. Stress in ASL: Empirical evidence and linguistic issues. *Language & Speech* 42: 229–250.

Wilbur, R. B. 2000. Phonological and prosodic layering of non-manuals in American Sign Language. In *The Signs of Language Revisited: Festschrift for Ursula Bellugi and Edward Klima*, H. Lane & K. Emmorey (eds), 213–241. Hillsdale NJ: Lawrence Erlbaum Associates.

Wilbur, R. B. 2003. Representations of telicity in ASL. *Chicago Linguistics Society* 39, 354–368.

Wilbur, R. B. 2009. Effects of varying rate of signing on ASL manual signs and non-manual markers. *Language & Speech* 52(2–3): 245–285.

Wilbur, R. B. 2010. The semantics-phonology interface. In *Sign languages*, D. Brentari (ed.), 357–382. Cambridge: CUP.

Wilbur, R. B. 2011a. Nonmanuals, semantic operators, domain marking, and the solution to two outstanding puzzles in ASL. *Sign Language & Linguistics* 14: 148–178.

Wilbur, R. B. 2011b. Sign syllables. In *The Blackwell Companion to Phonology*, M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice (eds), 1309–1334. Oxford: Blackwell.

Wilbur, R. B. & Patschke, C. 1998. Body leans and marking contrast in ASL. *Journal of Pragmatics* 30: 275–303.

Wilbur, R. B. & Malaia, E. 2008. Contributions of sign language research to gesture understanding: What can multimodal computational systems learn from sign language research. *International Journal of Semantic Computing* 2(1): 5–19.

Wilbur, R. B. & Schick, B. S. 1987. The effects of linguistic stress on ASL signs. *Language & Speech* 30: 301–323.

Language structure and principles of information organization

An analysis of retellings in Japanese, German, and L2 Japanese

Naoko Tomita

Heidelberg University

The empirical study investigates how native speakers of Japanese, native speakers of German and advanced German learners of Japanese organise information for establishing coherence when retelling a film story. There are differences found between the two L1 groups: (a) The speakers of German employ various lexical items (e.g., *dann* 'then') for organization of a text-intrinsic shift in the temporal domain, whereas the verbal aspect in Japanese suffices for this purpose. (b) The speakers of Japanese often take the protagonist's point of view for description of the situation which the protagonist perceives at his *now* in the narrative world, whereas the speakers of German adhere to their own perspective. (c) This is connected with the different ways for establishing causal coherence. Moreover, the analysis of the L2 retellings suggests that the principles of information organization in the target language are not evident to the adult learners mind and eye. These results are discussed in respect of the relationship between the typological features of the L1/L2 ([+/− verbal aspect, +/− subjectivity-prominent]) and the principles of information organization. In particular, the study focuses on the role of the grammaticised notions STATE, EVENT, and POINT OF VIEW in L1/L2 production.

1. Introduction

Information organization for the conceptual domains of TIME and ENTITY is the most relevant factor for establishing coherence in narrative texts, as text question "What happened at time *x* to protagonist *y*?" (von Stutterheim 1997) concisely illustrates.[1] Several cross-linguistic empirical studies in the last decade have shown that,

---

[1] For more details of the *quaestio*-approach to text production, see von Stutterheim & Carroll this volume.

in addition to this general principle, speakers follow a language-specific strategy for building up coherence. These studies have further provided evidence for the assumption that the language-specific preferences in information organization are (partly) induced by "selective attention" (Slobin 1996: 82) driven by the grammaticised conceptual categories. For example, an analysis of film retellings elicited from native speakers of aspect and non-aspect languages by Stutterheim and Nüse (2003) demonstrated that the typological feature [+ progressive/imperfective aspect] (e.g., English) facilitates speakers parsing event sequences into comparably small units, selecting different phases for encoding, and preferring a global perspective under which event times are hooked up to the *now* of the speaker as the 'experiencer' of the stimulus film ('deictic strategy', von Stutterheim et al. 2003: 109). In contrast, speakers of languages in which the aspectual notion of *ongoingness* is not grammaticised (e.g., German) encode events holistically, and this provides the relevant informational components for anaphoric shift in the temporal domain ('anaphoric strategy', von Stutterheim et al. 2003: 112).

In the field of SLA, it was also documented that principles of information organization related to grammaticised conceptual categories in the speaker's L1 are resistant to restructuring in adult second-language acquisition. For example, the *thinking-for-speaking* hypothesis proposed by Slobin (1996) on the basis of extensive cross-linguistic analyses of L2 and L1 narratives claims that the language-specific ways of perspective-taking which are influenced by grammaticised conceptual categories such as *temporal aspects*, *evidentiality*, and *voice* are difficult to be retrained, because those conceptual categories "are not categories of thought in general, but categories of thinking for speaking", and "cannot be experienced directly in our perceptual, sensorimotor, and practical dealings with the world" (Slobin 1996: 91). In contrast, L2 speakers would not have insurmountable difficulty in learning, for example, to mark the category of plurality in L2, because the category is evident to the non-linguistic mind and eye. In accordance with the *thinking-for-speaking* hypothesis, von Stutterheim et al. (2012) describe the 'nature' of grammaticised conceptual categories and their role for the L1 speakers in information organization as follows: Grammaticised categories are

> not only obligatory but highly abstract (i.e., applicable to word class paradigms, independent of the items' specific meaning) and are fully automatised in use. They provide a conceptual grid or frame for mental processing and decision making – at least whenever language is involved.   (von Stutterheim et al. 2012: 366)

The following two points summarise what the previous studies suggest in respect of language-specificity in principles of information organization in narratives:

(a) Grammaticised conceptual categories influence the strategy which native speakers use for building up coherence. This is also the case when speaking in L2.

(b) This relevance of grammaticised conceptual categories derives from the fact that they are abstract enough to be applied to diverse lexical contents, and therefore serve as a schematic, constant framework when constructing a 'preverbal message'[2] to be expressed.

Against this background, the aim of the present study is twofold: First, to illuminate how information is organised for establishing coherence in film retellings in L1/L2 Japanese and L1 German, and, second, to gain more insight into the role of grammaticised conceptual categories in establishment of coherence by L1/L2 speakers.

Many of the prior empirical studies on the relationship between grammaticised categories and principles of information organization have focused on the typological features [+/− verbal aspect], [+/− evidential inflections], [satellite-framed vs. verb-framed], [subject-first constraint vs. free word order] and [+/− assertion-related particles] [3] as independent variables, using appropriate test languages.[4] The present study not only revisits one of the typological oppositions just enumerated, i.e. [+/− verbal aspect], but also addresses the opposition [+/− subjectivity prominent] (see below). The opposition [+/− subjectivity prominent] has scarcely been covered in the previous investigations on the topic, but may possibly play a significant role in conceptualization processes in language production, and thus deserves an examination in respect of the *language-cognition* interface. Ikegami (2005, 2008, 2016) introduces the term 'subjectivity-prominent' for a typological characterization of

---

2. Psycholinguistic studies on language production generally assume that the speaker selects those information units which she/he wants to express and structures them under a specific perspective during planning processes of the 'preverbal message', which is then transformed into the one-dimensional medium language (Levelt 1989, 1999).

3. Dimroth et al. (2010) understand the genetic contrast between Germanic and Romance in relation to the following opposition between availability/non-availability of a special group of scope particles: Stressed variants of the particles *toch/wel* in Dutch and *doch/wohl/schon* in German (termed as *assertion-related particles*), which mark that the utterance in which they appear is in contrast to an earlier otherwise comparable utterance, lack a direct translation equivalent in Italian and French.

4. Test languages were: Aspect languages such as English and Spanish vs. non-aspect languages such as German and Hebrew (Slobin 1996; von Stutterheim & Nüse 2003); a non-evidential language, English, vs. an evidential language, Turkish (Slobin 1996); satellite-framed languages such as English, Mandarin, and Russian vs. verb-framed languages such as Spanish, Turkish, and Hebrew (Berman & Slobin 1994; Slobin 1996); English as a SVO language vs. German as a language with a relatively free word order (von Stutterheim & Carroll, this volume); and German and Dutch, with assertion-related particles vs. French and Italian, which lack assertion-related particles (Dimroth et al. 2010).

Japanese by reinterpreting Langacker's (1990: 17–21) dichotomy 'subjective construal' vs. 'objective construal'. According to Ikegami, Japanese (grammatically) allows zero encoding of the perceiver/conceptualiser of a situation to a greater extent than Indo-European languages in general; Japanese is a 'subjectivity-prominent' language. Sentence translations between English and Japanese in Ikegami's papers (2005, 2008, 2016) suggest that speakers of Japanese would possibly show a preference for such a zero encoding during speech production.

The present paper will first of all illustrate the linguistic categories in Japanese which constitute its typological features ([+ subjectivity-prominent] and [+ verbal aspect]) – that is, (a) grammatical markers of point of view which indicate subjective representation of a situation, and (b) the verbal aspect system. These grammatical categories lack equivalent counterparts in German. After that follows a text-linguistic analysis of film retellings, illuminating how native speakers of Japanese organise information for establishing coherence in contrast to native speakers of German. These results will be discussed in respect of the relationship between the typological features and the principles of information organization. This is complemented by a second analysis which compares information organization in retellings of the same stimulus film elicited from German adult learners of Japanese with retellings by native speakers of Japanese. This will give more insight into the persistency of the language-specific principles related to the typological features of the speaker's L1.

2.   Grammaticised notions in Japanese

With regard to the typological features of Japanese ([+ subjectivity prominent] and [+ verbal aspect]), the present section illustrates the grammatical markers of point of view and the verbal aspect system in the language. These grammatical categories embody a particular abstract notion, i.e., a particular perspective which can be taken for diverse lexical contents. They therefore may serve as a schematic, constant framework for speakers.

2.1   Markers of the point of view

The point of view from which the speaker describes a situation can be manifested in different linguistic forms. The examples which best illustrate this both in German and Japanese are deictic words, which can be used if the speaker's *I-here-now* constitutes the reference point of a situation (e.g., *hier* and *koko*, referring to a place near the speaker vs. *dort* and *asoko*, referring to a place away from the speaker, or *jetzt* and *ima,* referring to the time of the speaker's utterance). In addition to such lexical

words, Japanese has grammatical forms available for indicating the speaker's orientation. Previous studies have enumerated 'auxiliary verbs'[5] like *-teku-(ru)* 'toward here', *-teik-(u)* 'away from here', *-teyar-(u)*, [6] *-tekure-(ru)*, and *-temora-(u)* (Okugawa 2007: 33; Tomita 2008: 216–221; 280; Narita 2009: 404; Takemura 2010: 291; Wei 2010: 259). In what follows, the way these 'auxiliary verbs' indicate the point of view taken by the actual speaker will be exemplified on the basis of *-teku-(ru)*.[7]

The meaning of *-teku-(ru)* traces back to the meaning of its origin – the lexical (full) verb *ku-(ru)* 'to come'. Kuno (1978) describes the relationship between the spatial configuration of translational motion on the real or narrative-fictive world – this is physically 'given' – and the point of view the actual speaker takes for expression of the motion, as follows:

> [The full verb][8] *ku-(ru)* is used if the point of view taken by the speaker is approximate to the position of someone whom the moving entity approaches.
>
> (Kuno 1978: 253, translated by author)

Referring to this, I assume that the use of the 'auxiliary verb' *-teku-(ru)* in narrative texts indicates that the speaker/narrator takes over the point of view of a figure approached by a moving entity. For example, clause 64 in (1) (see below) refers to an event that involves translational motion, i.e., the coming out/towering up of a big stone block. *-Teku-(ru)* as used in the clause (in the form *-teki-*) adds the information that the coming out of the stone block was directed "to me, the speaker"; it was "at my feet". Actually, the stone is not towering in the direction toward the narrator at the time of ongoing narration, but toward the protagonist at his *now*. Thus, the use of *-teku-(ru)* indicates that the speaker/narrator takes the point of view of the protagonist and describes what situation the protagonist perceives at the protagonist's *now* and present position.

---

5. In the present paper, the term 'auxiliary verb' (in single quotes) is used for the term *hojo-dooshi* in Japanese linguistics. *Hojo-dooshi* constitutes a predicate, so-called verbal complex, when joined to their main verb, where its original lexical meaning is not fully bleached, in contrast to delexicalised English auxiliary verbs such as *to have* in perfect construction or *to be* in passive construction.

6. The 'auxiliary verbs' *-teyar-(u)*, *−tekure-(ru)*, and *-temora-(u)* stem from three lexical (full) verbs for giving and receiving. Depending on the speaker's point of view, one of these lexical verbs is chosen for expression of a giving/receiving event. The related 'auxiliary verbs' can be used to indicate the point of view taken even when the event to be described does not involve a concrete action of giving/receiving.

7. Of these 'auxiliary verbs' for marking the point of view, *−teku-(ru)* is most frequently used in the retellings which were collected for the analysis in the present paper.

8. Added by author.

(1)   jj08: 63–64
      63   *arui-teiku-to*[9]
           walk-PROG:N-PAST-CONJ
           '(The protagonist) is walking thither, then'
      64   *totsuzen   ookina   ishi-ga   de-teki-te*
           suddenly   big   stone-NOM   come out-POV-CONJ
           'a big stone block suddenly towers up (toward me, the narrator), and […]'

By taking the protagonist's point of view, the narrator's *I-here-now* overlaps with the protagonist's *I-here-now* in this example. Thus, the events are told from inside the narrative world, which creates vividness for narration.

Through grammaticalization processes, the meaning of the origin – the lexical (full) verb *ku-(ru)* 'to come' – is bleached to the extent that the derived 'auxiliary verb' *-teku-(ru)* can also be used for expression of a situation which does not involve directed motion, but implies someone perceiving it. For example, the usage of *-teku-(ru)* in clause 13 in (2) (see below) adds the information that it is "me, the speaker" who is hearing some noises, and so indicates that the speaker/narrator takes the point of view of the protagonist. The narrator's *I-here-now* overlaps with the protagonist's *I-here-now* in this example.

(2)   jj08: 12–16
      12   *mawari-o   miru-to*
           surroundings-ACC   look:N-PAST-CONJ
           '(The protagonist) looks at surroundings, then'
      13   *nanika   oto-ga   shi-teki-te*
           some   noise-NOM   sound-POV-CONJ
           'some noise sounds (to me, the narrator), and'
      14   *sora-o   miageru-ga*
           sky-ACC   look up:N-PAST-but
           '(he/I) look(s) up to the sky, but'
      15   *ame-wa   fut-teko-nai-node*
           raindrops-TOP   fall-POV-NEG:PRS-because
           'raindrops do not fall (toward me, the narrator), for this reason'
      16   *shita-o   mi-temiru.*
           down-ACC   look-TEN:N-PAST
           '(he/I) looks down tentatively.'

Organizing the point of view through the use of an adequate 'auxiliary verb' relieves the need to establish causal coherence. For example, the action of the protagonist

---

9.   The bounded temporal conjunction *-to* specifies the time for which the situation described in the subsequent clause holds (for more details, see Tomita 2008: 154–157). This conjunction roughly means 'exactly at this point in time'. For convenience, it is translated with *then* in the present paper.

expressed in 14 in (2) can be understood as a reaction to the auditive perception as cause, because the protagonist alias the narrator is involved as the perceiving person in situation 13 in (2).[10] In contrast, the narrator of the following German Example (3) has difficulties herewith because of the use of the indefinite pronoun *man* 'one' to refer to the perceiving entity.

(3) gg12: 13–14

13 *man hört   dann langsam   n tropfen also wassergeräusche also*
one   hear:PRS then gradually a drop   well water sounds   well
so'n   leichtes tropfen im hintergrund.
such a quiet   drip   in   the background
'Then, one gradually hears a drop, well, water sounds, well, such a quiet drip in the background.'

14 *das männchen scheint   draufhin   was   zu suchen.*
the little man   seem:PRS thereupon something to seek
'The little man seems thereupon to seek something.'

It is not the aim of the present paper to describe all of the 'auxiliary verbs' in the ways in which they function as markers of the point of view taken by the actual speaker/conceptualiser. Nonetheless, Japanese deserves to be typologically classified as 'subjectivity-prominent' (not only, but above all) because of the availability of a series of those grammatical devices which allow zero encoding of the speaker as the perceiver who observes/conceptualises the event to be expressed from a particular view point. It is the notion POINT OF VIEW which is grammaticised in Japanese, in contrast to German.

2.2   Aspect system

In line with a general assumption in Japanese linguistics (Okuda 1985a, 1985b; Kudo 1995), I take the position that the two verbal forms in Japanese, i.e., the simple form and the *tei*-form, indicate particular complementary aspectual notions. In this section, I will illustrate these aspectual notions, using the aspectual labels *non-imperfective* (the simple form) and *imperfective*[11]/*resultative* (the *tei*-form). After that, a brief look at the temporal meaning of German verbs in the tense

---

10. The same applies to the relationship between the action presented in 16 in Example (2) and the situation 15 which is perceived by the protagonist alias the narrator.

11. In previous literature, the label *continuous* ('*doosa no keizoku*') (Okuda 1985a, 1985b; Kudo 1995) is used to refer to the imperfective meaning of the *tei*-form. In the present paper, the term *imperfective* is preferred, since the form can also be used for habitual events (e.g., *Kare-wa maiasa koohii-o non-dei-ru* 'He drinks coffee every morning'), or for a generic description (e.g., *Chikyuu-wa 24-jikan-ni ikkai jitenshi-tei-ru* 'The Earth rotates on its axis once every 24 hours').

marked simple form and the English aspect system will provide more insight into the properties of the Japanese aspect system.

The *tei*-form, regardless of whether it is used in the imperfective or in the resultative meaning, presents a situation as a STATE without any boundaries, where the term *boundary* refers to points in time such as *starting point*, *endpoint*, and *transition point to a post-state* concerning a described situation. Sentence (4) illustrates this for the imperfective *tei*-form:

(4) *Ferikkusu-ga warat-tei-ta.*
     Felix-NOM   laugh-IMP-PAST
     'Felix was laughing.'

The aspectual relation expressed by the *tei*-form in (4) defocuses both the starting point and the endpoint of the situation of *laughing of Felix*. To put it another way: the *tei*-form marks the subinterval of the respective situation as selected for the *topic time* (i.e., the time about which an assertion is made; Klein 1994).

Situations such as *tidying up the room* and *crossing over a wide river* have in common that they involve a quality change either of the entity in the object function or the entity in the subject function. For this type of situation, there are two phases which the aspectual relation of the *tei*-form can apply to, so that the *tei*-form can be interpreted either as imperfective or resultative:

(5) *Ferikkusu-ga heya-o   katazuke-tei-ta.*
     Felix-NOM   room-ACC tidy up-IMP-PAST   (the imperfective reading)
     'Felix was tidying (up) the room.'
     Felix-NOM   room-ACC tidy up-RSLTV-PAST   (the resultative reading)
     'Felix had tidied up the room (and the room was tidy).'

The resultative relation expressed by the *tei*-form in (5) defocuses both of the following points in time: First, the transition point to a post-state, which is also the starting point of the post-state of *tidying up*, and second, the endpoint of the post-state of *tidying up*. In other words: The resultative *tei*-form marks the subinterval of the post-state as selected for the topic time. This means that, for a resultative sentence, the topic time corresponds to a stative phase, instead of to the dynamic phase of the situation during which the entity in the subject function is acting in some way.

As this illustration has demonstrated, both aspectual relations of the *tei*-form (i.e., *imperfective* and *resultative*) share the following feature: All of the points at which a change occurs (i.e., the *starting point*, the *endpoint*, the *transition point to a post-state*, and the *endpoint of the post-state*) are defocused. This temporal semantic feature is of central importance for textual functions, since the situations presented under the perspective expressed by the *tei*-form cannot trigger an anaphoric shift in time. As an approximation, consider the simultaneousness in the sentence pair *Dorothy was putting her golden cap on and was reciting the magic poem*. In the

present paper, the perspective marked by using the *tei*-form is termed STATE, regardless of whether the *tei*-form marks an imperfective relation or a resultative relation.

In contrast to the *tei*-form, the simple form presents a situation as EVENT with at least one boundary, that is, a point in time such as a *starting point*, *endpoint*, or *transition point to a post-state* concerning a described situation. The following will illustrate this: For situations without an inherent state change (e.g., *laughing of Felix*, see [6]), there are two alternative interpretations with respect to the relation between the topic time and the time of situation; the topic time either includes only the starting point of the situation of *laughing*, or it includes both the starting point and the endpoint.

(6) *Ferikkusu-ga warat-ta.*
    Felix-NOM   laugh-PAST
    'Felix laughed.'

Due to the fact that the former reading is as likely as the latter reading, the term *non-imperfective* is more suitable than the label *perfective* for the simple form in Japanese. When applied to change-of-state situations (e.g., *opening of a window* or *crossing over a wide river*), the simple form expresses that the topic time includes the transition point to a post-state. Altogether, with regard to situations with and without an inherent state change, all of the possible aspectual readings of the simple form share the common feature that the topic time includes at least one boundary. Based on this temporal semantic feature of the simple form, situations presented by using this form trigger an anaphoric shift in time. This corresponds approximately to the successiveness in the sentence pair *Dorothy put her golden cap on and recited the magic poem*. In the present work, the perspective expressed by using the simple form is termed EVENTS, regardless of which point in time is included in the topic time.

In contrast to Japanese, in which the complementary aspectual notions STATE (without any boundaries) and EVENT (at least one boundary) are grammaticised, German has no grammatical aspect for *ongoingness*, and tense marked (simple) forms in German in themselves are ambiguous with respect to the assertion of the endpoint of the situation described. A clear interpretation can be derived from empirical knowledge shared by the speaker and hearer. Compare hereto the finite verb *wickelte* 'swaddled' in the following two different contexts: *Eine Mutter wickelte ein Baby in einen Leinentuch, als ich das Zimmer betrat* 'A mother was swaddling a baby in linen as I entered the room' and *Eine Mutter wickelte ein Baby in einen Leinentuch und ging aus* 'A mother swaddled a baby in linen and went out'.[12]

---

12. Bohnemeyer and Swift (2004: 269) give further examples which show that a German simple form used for expression of a (inherently) telic situation is preferably related to a 'perfective' meaning, but this implicature can easily be cancelled in an appropriate context.

In another aspect language, English, the aspectual notion *progressive* is grammaticised. With this, it would appear on a first glance that Japanese and English could be brought together under one language-typological category.[13] However, the two aspect systems – the Japanese with the *tei*-form and the English with the *ing*-form – differ from each other with respect to the extent to which the defocused *endpoint* (or *transition point)* still plays a role in the background of the speaker's conceptualization processes. For perspective-taking with the *tei*-form, regardless of whether the *imperfective* or *resultative* perspective is meant, it does not matter whether the situation to be described (e.g., *laughing, standing in front of the whiteboard, or knowing*) is dynamic and involves an inherent endpoint. Instead, the starting point of the situation described by the descriptive content plays a relevant role for the STATE perspective expressed by the *tei*-form.[14] In contrast, the English progressive can be defined as "the combination of progressive meaning and nonstative meaning" (Comrie 1976: 35, and many other authors). In other words, different from the *tei*-form, the use of the *ing*-form requires an inherent endpoint of the situation to be described. Moreover, by using the *ing*-form, this situation is presented as a "contingent and make-believe state" (Comrie 1976: fn. 1 on page 49). Thus, the defocused *endpoint* still exists in the background of the speaker's mind, as Jesperson (1924) already described just under a century ago:

> It is a natural consequence of the use of the expanded tenses [i.e., *ing*-form][15] to form a time-frame around something else that they often denote a transitory as contrasted with a permanent state which for its expression requires the corresponding unexpanded tense. The expanded form makes us think of the time-limits, within which something happens, while the simple form indicates no time-limit.
>
> (Jesperson 1924: 279)

---

13. Sugaya and Shirai (2007) seem to hold this view, claiming that "the Japanese tense-aspect system has much in common with that of English. […] Similar to the English progressive form *be -ing*, Japanese has an aspect marker *-te i-(ru)* that must be used in describing action in progress at the time of reference" (Sugaya & Shirai 2007: 4–5). The authors use thus the term *progressive* for the imperfective meaning of the *tei*-form. With respect to the role of the *endpoint* (see below), however, it is questionable whether the term *progressive* is the most adequate label for the *tei*-form. For the choice of the label 'imperfective' in this paper, see also fn. 11.

14. Teramura (1984: 127) claims hereto that "the Japanese *continuous* (/*imperfective*) form expresses that the result of a beginning now exists" (translated by author). See also the interpretation of the *tei*-perspective by using the concept of *reference point* for an anaphoric shift in time in Tomita (2008: 91).

15. Added by author.

The empirical analysis of narrative texts (Section 3) will shed light on whether this difference between the Japanese and English aspect systems materially impacts how information is organised in texts.

Up to this point, we have considered conceptual categories which are grammaticised in Japanese and constitute the two typological oppositions [+/− verbal aspect] and [+/− subjectivity prominent] between Japanese and German. Due to the availability of these grammaticised categories, speakers of Japanese are faced with the following two kinds of decision-making each time when they construct a 'preverbal message' (see fn. 1): the choice between the STATE perspective and the EVENT perspective, and the choice between the speaker's point of view (i.e., the default perspective) and the point of view of someone other than the speaker.[16] This obligatory decision-making leads speakers of Japanese, in contrast to speakers of German, to attend to the notions STATE, EVENT, and POINT OF VIEW, when generating a "preverbal message".

Thus, based on findings from previous studies suggesting a relationship between grammaticised categories and principles of information organization (see introduction), we can hypothesise for the present study that the typological differences between Japanese and German ([+/− verbal aspect] and [+/− subjectivity prominent]) entail that both speaker groups prefer different ways of information organization for retelling the same episode. The subsequent section presents an empirical analysis which aims to illuminate how information is organised for building up coherence in film retellings by native speakers of Japanese in contrast to by native speakers German. The results will be discussed with respect to the typological differences described above.

3.   Information organization in Japanese narratives in contrast to German narratives

The textlinguistic analysis in the present section is based on retellings of the silent film "Quest" by Thomas Stellmach (four texts in Japanese and three texts in German). In this seven-minute movie, the protagonist, a clay man, travels through

---

16. The second type of decision, i.e., whether or not to take a particular, marked perspective, is not limited to categories which are grammaticised in the language. It can be assumed that the speaker has to decide between *he is tired* and *he is probably tired* whenever she/he wants to produce an utterance in a language which has the lexical word *probably* available. Thus, not only grammatical categories but also lexical items could influence patterns in information organization, if they convey a conceptual category (in this case, the *epistemic value*) which can serve as a schematic framework for the speaker when constructing a 'preverbal message'. The present study, however, does not address this aspect of particular lexical items.

five worlds (a desert world, a paper world, a stone world, and two worlds of industrial machines) in quest for water, and faces a series of obstacles raised by the environment. For illustration, a screenshot is cut from the film (Figure 1). The screenshot presents the protagonist hoping that water will fall from the sky.

Figure 1. The clay man in the paper world

Retellings were collected from adult native speakers of Japanese and of German using the following method: Participants first viewed the stimulus in full. They then viewed the film again, this time stopping at the end of each world the protagonist walked through. Each speaker was asked to describe what happened in the world she/he had just seen. The corpus studied consists of the retellings of the episode in the first three worlds. The average length for the Japanese group is 84 clauses, for the German corpus it is 76 clauses.

The text analysis in what follows is based on a quantitative and qualitative analysis by Tomita (2008).[17] This German-Japanese contrastive analysis addresses two central aspects of macro-structural information organization in narratives: a) The global temporal structure, and b) the role of the global topic entity (in this case, the sole protagonist) to establish coherence. In this section, principal findings from the monograph are reformulated in respect of the present study's concerns.

3.1   The global temporal structure

Both Japanese and German speakers prefer to represent the events in the narrative world which constitute the storyline as a temporally bounded situation which thus provides the anchor point for an anaphoric, text-intrinsic shift in the temporal

---

17. Numbers of occurrences of relevant linguistic forms in the retellings, which then are interpreted in respect of the lexical content of the linguistic context, are reported in tables 5.1–5.17 (for L1 Japanese) and Tables 6.1–6.10 (for L1 German) in Tomita (2008).

domain (termed 'anaphoric strategy' by von Stutterheim et al. 2003: 112). The two speaker groups use different linguistic means for this type of temporal organization, however. The speakers of German employ various lexical items, whereas the verbal aspect in Japanese suffices for this purpose. The following German example illustrates that anaphoric shift in time is explicitly marked by the temporal *shift*-adverb *dann* 'then' (see clauses 64, 69, 73, and 75).

(7)   gg06: 63–75
   63   *und schaut sich um in dieser öden welt*
       'and (he) looks around in this dead world'
   64   *und sieht dann plötzlich auch da unten so'n stein*
       'and then (he) sees suddenly a stone down there'
   65   *wo so'n felsbrocken richtig/*
       'where a certain rock (is) really …'
   66   *wo wasser drauf is*
       where there's water on it.
   67   *so'n n bisschen/*
       'a little bit …'
   68   *der eben nass is*
       'which is namely wet'
   69   *und dann geht er dazu/dahin*
       'and then he goes thither, to there'
   70   *und fasst diesen stein an*
       'and takes this stone'
   71   *und sieht*
       'and looks'
   72   *dass er eben feucht ist*
       'that it is just wet'
   73   *und vermutet dann natürlich*
       'and then (he) expects, of course'
   74   *dass darunter wahrscheinlich dann endlich das erhoffte wasser ist*
       'that, probably, there is finally the water wished for'
   75   *nimmt dann sich einen anderen stein zu hilfe.*
       'and then (he) takes another stone as an aid.'

This example further demonstrates that situations which are marked as a delimited situation with the use of lexical words induce the interpretation that a shift in time occurs (see the temporal adverb *plötzlich* 'suddenly' in 64, the verb particle *dahin* 'thither, to there' in 69, and the change-of-state verbs *anfassen* 'to take hold' in 70 and *sich nehmen* 'to take' in 75).

In the Japanese retellings, in contrast, situations presented under the EVENT perspective by using a verbal aspect form (i.e., the simple form) trigger an anaphoric shift in time. Use of lexical *shift* markers such as *sono-ato/sorede/soshite* 'and

(then)' is marginal (12 times in four retellings), as the following example illustrates (the simple form is marked in bold).

(8) jj03: 54–60
    54 *ashi-o **kake**-te*[18]
        '(He) puts his foot on (the rock), and'
    55 *iwa-ga **kake**-te*
        'the edge of the rock comes off, and'
    56 *kare-wa ochi-**teshimaimasu.***
        'he falls down.'
    57 *yatto mizu-o **mitsuke**-te*
        'Finally, (he) finds water, and'
    58 ***yorokon**-de*
        '(he) rejoices, and'
    59 *horeru-kana-to **kangae**-te*
        '(he) thinks that one could dig, and'
    60 *togatta iwa-de hori-**hajimemasu.***
        '(he) starts to dig with a spiky stone.'

Another point that characterises the Japanese type of 'anaphoric strategy' is the way in which situations presented under the STATE perspective (e.g., 'the man was looking down carefully' (jj08: 71) or 'water was dropping' [jj04: 51]) are involved in the linear narrative discourse. Due to the lack of any temporal boundary, a situation presented under the STATE perspective by itself does not trigger an anaphoric shift in time (see Section 2.2); actually, it violates the strategy of constant time shifts in narrative texts. Our data show that this undesired effect of a STATE is unfailingly reconciled by encoding the situation in a subordinate clause,[19] or by marking the point of view of the global topic entity. Let us examine the situation of *dropping of water* which is expressed under the STATE perspective with the *tei*-form in the main clause 51 in Example (9).

(9) jj04: 50–51
    50 *mi-temiru-to*
        '(he) looks around tentatively, then'
    51 *so-no-ashimoto-ni nanto mizu-ga potapotato tare-tei-mashita.*
        'indeed, water was dropping at the feet, with plopping sounds.'

---

18. If a verb is followed by the connective particle *-te*, the tense morpheme is omitted, whereas the aspectual opposition can still be marked by the simple form (i.e., V-te) vs. the *tei*-form (i.e., V-tei-te).

19. Through the choice of the type of subordination, the speakers of Japanese differentiate between STATES in terms of the topicality of the entity referred to. See Tomita (2008) for more details.

The marker of tentativeness *-temi-(ru)* in the precedent clause 50 and the modal particle *nanto* 'indeed' in clause 51 are revealing. The tentative marker expresses that the protagonist looks around in the expectation of finding an answer. The particle *nanto* used in 51 marks that the situation of dropping of water was found by the person who is 'now' looking for an answer. Altogether, the interpretation suggests itself that the speaker/narrator takes the point of view of the protagonist, who is looking around for water, and describes what situation is perceived at this particular point in time from that particular standpoint. Thus, this situation, which is expressed under the STATE perspective, can fill a slot on a time line of the narrative text.

3.2   The role of the global topic entity in information organization

There are numerous studies which compare contexts in which the topic particle *-wa* in Japanese is attached to the subject NP with contexts in which only the nominative marker *-ga* is used for the subject NP. For theoretical discussions about functions of *-wa* in this line, see Chafe (1976), Lambrecht (1994), Fujinawa (2017), and many others. For empirical analyses of the use of *-wa* in narrative discourses, cf. Clancy & Downing (1987) and Maynard (1987). The present study is rather interested in examining the role of the topic entity (i.e., the sole protagonist in the film story) in construing situations, regardless whether it is overtly expressed in the sentence, or not.

The analysis revealed for both speaker groups that the global topic entity, i.e., the sole protagonist, plays a significant role in establishment of coherence, serving as the constant topical element in the text. For the German data, the protagonist is involved as a verb argument in 90.7% of the clauses which express situations constituting the storyline; this is the case in 82.5% of the Japanese data. In both data sets, this argument functions almost exclusively as the subject of the clause, and is often realised as a zero form (see above, German Example (7) and for Japanese, Example [8]).

With respect to the organization of the point of view from which the story is narrated, however, there is a considerable difference between the narratives in the two languages. The speakers of Japanese often conceptually include the global topic entity as a perceiving individual when describing a situation in which it is not involved as a (verb) argument (e.g., *towering up of a big stone block* or *dropping of water*). In other words, the speaker/narrator temporarily takes the protagonist's point of view for description of the situation which the protagonist perceives at his *now* and at his present position in the narrative world. The shift of the origo from the viewpoint of the speaker/narrator to the viewpoint of the protagonist is organised with specific linguistic means such as the spatial-deictic 'auxiliary verb' *-teku-(ru)* (as in clause 64 above in (1) and clauses 13/15 in [2]), the marker of

tentativeness *-temi-(ru)* (as in clause 16 in (2) and clause 50 in [9]), and the bounded temporal conjunction *-to*[20] (as in clause 63 above in (1), clause 12 in (2), and clause 50 in [9]), while a return to the default perspective does not require any specific linguistic indicators.

The constant shifting of the origo from the view point of the speaker/narrator to the view point of the protagonist, and vice versa, is one of the principles in information organization for narrative texts in Japanese, whereas the speakers of German adhere to 'one' perspective. For example, the indefinite pronoun *man* 'one' used in the introductory clause 28 in the German Example (10) (see below) is revealing, since the pronoun indicates that events on two levels are described from the speaker/narrator's viewpoint. Thereby, the first level is related to the actual context of the speaker in which the speaker or someone else perceives fictitious events occurring at the second level, i.e., in the narrative world presented in the stimulus video. Thus, situations referring to the protagonist's circumstances (i.e., *slipping back, sliding down through a hole*, and *disappearing of sand*) as well as a situation involving the protagonist (i.e., *drawing the protagonist into the hole*) are constantly narrated from the speaker/narrator's viewpoint. Such a consistency of the perspective in the German retellings can also be seen in that a third person pronoun (*er, dieser*) is constantly used for an overt reference to the protagonist.

(10)   gg03: 28–33
28   *und plötzlich sieht man*
     'and suddenly one sees'
29   *der sand gibt nach*
     'the sand slips back'
30   *scheint so durch ein loch unter der oberfläche hindurch zu rutschen*
     '(the sand) seems to slide down through a hole under the ground's surface'
31   *wie in einer sanduhr*
     'like within a sand glass'
32   *und der sand äh verschwindet*
     'and the sand disappears'
33   *und mit dem sand wird das männchen mit in dieses loch hineingezogen*
     'and together with the sand, the little man is drawn into this hole.'

The difference in the way of organization of the point of view between the two speaker groups is accompanied by the different ways for establishing causal coherence. In the Japanese retellings (see (2) above), a situation referring to circumstances (i.e., *sounding of some noise*) and an action of the protagonist (i.e., *looking up to the sky*) can form a causal relation of the type 'psychological cause – reaction',

---

20. The conjunction roughly means 'exactly at this point in time' (see fn. 9).

since the protagonist's point of view is marked. When this point of view is indicated, the speaker/narrator needs only to assert the situation which motivated the protagonist – alias the actual speaker/narrator – to look up to the sky. In the German retellings, in contrast, anaphoric shift relations in time are primarily marked, which then provide the basis for causal coherence, since temporal shift relations can suggest a causal relation. For example, clauses 64–68 in (7) above refer to an event of visual perception by the protagonist, and the usage of the shift adverb *dann* 'then' in 69 explicitly marks that the actions of the protagonist (i.e., *going thither* and *taking the stone*) temporally follow this event of perception. This successive relation between the event of perception and the actions motivates the interpretation that the events are integrated in a causal relationship.

3.3   Discussion

The empirical analysis in the preceding section has shown that the two speaker groups have several common principles for establishing coherence: First, both follow an 'anaphoric strategy' for temporal organization; fictitious events are presented as a temporally bounded situation which provides the anchor point for a text-intrinsic shift in the temporal domain. Second, they integrate the global topic entity, that is, the sole protagonist, as the constant topical element in the text; the protagonist is involved in more than 80% of the clauses expressing a fictitious event, and is the subject of the clause in nearly all of these cases. In this way, the global topic entity contributes to establishing static coherence.

We must also, however, note the following differences between the German and Japanese ways of organizing information:

a. For the German 'anaphoric strategy', lexical words indicating a bounded view on the situation (such as the verb particle *dahin* 'thither, to there' and change-of-state verbs) or marking *shift*-in-time (such as *dann* 'then') play a significant role. In contrast, the grammatically encoded EVENT perspective on a situation suffices for anaphoric shift in Japanese narratives. Use of lexical *shift* markers in the Japanese data is marginal.

b. Speakers of Japanese have a keen sense of the undesired effect of the STATE perspective on the strategy of constant anaphoric shift in narrative texts. This effect is unfailingly reconciled by encoding the respective situation in a subordinate clause or by marking the point of view of the global topic entity.

c. A third difference between the two speaker groups concerns the organization of the point of view. The speakers of German adhere to 'one' perspective; they remain at their point of view, regardless of whether or not the situation described involves the protagonist as a (verb) argument. In contrast, one of the principles

in information organization in Japanese is the constant shifting of the origo from the view point of the speaker/narrator to the view point of the protagonist, and vice versa. Thus, the global topic entity is always present, whether included as an argument or not, as a central component of the conceptual structure.

d.   The Japanese way of organizing the point of view fosters that a causal linkage of the type 'psychological cause – reaction' is conceptualised between a situation referring to circumstances (i.e., *sounding of some noise*) and an action of the protagonist (i.e*., looking up to the sky*). In the German retellings, in contrast, anaphoric shift relations in time are explicitly marked, which then provides the basis for causal coherence.

In what follows, these differences in the principles in information organization between the two speaker groups will be discussed with regard to the availability of the notions STATE, EVENT, and POINT OF VIEW in the languages (see Section 2):

Ad (a) In contrast to Japanese, in which the complementary aspectual notions STATE (without any boundaries) and EVENT (at least one boundary) are grammaticised, German has no grammatical aspect for *ongoingness*, and tense marked (simple) forms in German in themselves are ambiguous with respect to the assertion of the endpoint of the situation described. Given this typological background, the empirical result (a) is in line with the following claim made in prior studies concerning conceptualization processes by speakers of German (von Stutterheim & Nüse 2003; von Stutterheim et al. 2003; von Stutterheim & Carroll 2006): The 'holistic view' (i.e., the assertion of the endpoint or the transition point to a post-state, von Stutterheim et al. 2003: 113) is compatible with linguistic means provided in the language and enables speakers of German to present 'reportable content' (von Stutterheim & Carroll 2006: 41). Thus, the notion of HOLISTICNESS is more salient than the notion of *ongoingness* for speakers of German during conceptualization processes, and this is the ground for which the German 'anaphoric' strategy (which was used in the corpus) is based on the 'holistic view', which was marked by various lexical words.

Ad (a) and (b) Speakers of Japanese follow a different 'anaphoric' strategy than speakers of German, i.e., the strategy which is based on the EVENT perspective expressed by an aspect form of the verb (the simple form). Together with the fact that the undesired effect of the STATE perspective (the *tei*-form) is unfailingly reconciled applying specific manners, this empirical finding means that the two grammaticised concepts EVENT and STATE play a prominent role in information organization; the Japanese way of temporal organization in narratives is based on the complementarity of these two aspectual notions.

Ad (c) and (d) The key fact here is that the availability of the markers of the point of view in the language does not necessarily mean that the speaker must use them in a particular context. Speakers of Japanese are not bound to shift the origo when retelling, but prefer to do so as a way to establish coherence. This preference is not shared by the speakers of German, in which the category POINT OF VIEW is not grammaticised.

In summary, the availability/non-availability of the notions STATE, EVENT, and POINT OF VIEW in the language influences the way information is organised by speakers of German and Japanese. This finding supports the present study's hypothesis: The typological differences between Japanese and German ([+/− verbal aspect] and [+/− subjectivity prominent]) entail that both speaker groups prefer different ways of information organization for retelling the same episode. At this point, a glance at another aspect language, English, is revealing. Previous studies (von Stutterheim & Nüse 2003; von Stutterheim et al. 2003; von Stutterheim & Carroll 2006) have shown that speakers of English prefer a 'deictic strategy' in which the speaker's *now* functions as the anchor point for event times (see introduction). The authors concluded that "the structural feature [+/− aspect] induces a specific pattern of event construal" (von Stutterheim & Nüse 2003: 870). However, the criterion [+/− verbal aspect] cannot explain why speakers of Japanese and English follow different strategies for temporal organization, since, as is well-known, the two languages share the common feature [+ verbal aspect] (as well as [+ progressive/imperfective aspect]). Only a detailed analysis sheds light on the conceptual differences between the two aspect systems,[21] and only by regarding those conceptual differences can we understand that the 'deictic strategy' exactly matches the conceptual properties of the progressive perspective marked by the *ing*-form. The same applies for the relationship between the Japanese 'anaphoric strategy' and the conceptual properties of the Japanese aspect system with the two complementary concepts EVENT and STATE. On these grounds, I propose that a fine-grained analysis of language-specifically grammaticised notions should form a basis for further investigation of the relationship between the language structure and the principles of information organization.

---

21. As discussed in Section 2.2, the notion of STATE (the *tei*-fom) is dissimilar to the perspective marked by the English *ing*-form in respect of the role of the defocused *endpoint* of the situation.

4. Information organization by German adult learners of Japanese

This section presents an empirical analysis comparing information organization in retellings by German adult learners of Japanese with those by native speakers of Japanese. Previous studies (the *thinking-for-speaking* hypothesis, see introduction) have shown that language-specific principles which are related to grammaticised conceptual categories in the speaker's L1 are resistant to restructuring in adult second-language acquisition. By including L2 Japanese data into this research context, I aim to obtain more insight into the persistency of those language-specific principles.

4.1   Data collection

This analysis is based on retellings of the silent film "Quest" collected from 24 German mid- to advanced-level learners of Japanese (6 female, 18 male, 21–49 years old). The same method of data collection was used as for the L1 corpus (see Section 3). The four L1 Japanese retellings are used as reference data for the present analysis of the learner corpus. All of the L2 speakers learned Japanese in a classroom setting and successfully acquired a certificate conforming to approx. 600 hours of learning (e.g., JLPT, level N2). Most of them were studying Japanese studies at the University of Heidelberg at the time of data collection. The average length of the retellings in L1 Japanese is 84 clauses, for the retellings in L2, it is 42.5 clauses (336 propositions in the L1 Japanese data and 1019 propositions in the L2 Japanese data).

4.2   Method

Regarding the differences in information organization between the two related native-speaker groups (see Section 3.3), the following four frequencies in each corpus were used to examine the extent to which the L2 speakers reorganise the language-specific principles in information organization:

a. The frequency of inadequate use of the *tei*-form (STATE perspective) for situations in which the protagonist is not involved as a (verb) argument (e.g., *dropping of water*);
b. The frequency of lexical words which mark an anaphoric shift in the temporal domain (e.g., *sono-ato* 'thereupon') or implicate it (e.g., *sorede/soshite/sore-kara* 'and (then)');
c. The frequency of a linguistically marked shift of origo from the point of view of the speaker/narrator to the point of view of the protagonist; and
d. The frequency of linguistically marked epistemic states of the speaker/narrator concerning thoughts and emotions of the protagonist.

The lower the frequency (a) and (b) are in the L2 corpus, the more target-language-like the temporal organization in the L2, since, as found in Section 3, the undesired effect of the STATE perspective on the strategy of constant anaphoric shift is unfailingly reconciled by the L1 speakers by using particular linguistic means (i.e., subordination and the markers of the point of view). Furthermore, the simple form (EVENT perspective) suffices for anaphoric shift. Frequencies (c) and (d) concern the organization of the point of view. A high frequency of marked shift of origo and a lower frequency of marked epistemic states of the speaker/narrator suggest target-language-like information organization, since the L1 speakers of Japanese take the point of view of the protagonist to describe a situation perceived by the protagonist, as well as for expression of thoughts and emotions of the protagonist. This 'direct' way of presenting the protagonist's inner world does not allow an expression of the speaker's/narrator's epistemic states.[22]

4.3   Results

Occurrences of critical expressions (a) – (d) (see Section 4.2) are dissimilar between the L1 and L2 corpus.[23] The numbers in Table 1 suggest that the L2 speakers were guided rather by the principles for information organization in their L1.

Table 1. Occurrences of critical linguistic expressions in the L1 and L2 Japanese corpus

	(a) inadequate use of the *tei*-form	(b) lexical words for temporal shift	(c) shift of origo	(d) epistemic states of the speaker/ narrator
L1 4 speakers (336 propositions)	0	12	21	0
per 100 propositions	0.0	3.57	6.25	0.0
L2 24 speakers (1019 propositions)	73	84	9	18
per 100 propositions	7.16	8.24	0.88	1.77

---

22. See also examples in Zeman, this volume, which illustrate that, on the strength of their deictic structure, epistemic modalities may function as an indicator of a divergence between the narrator and the character level.

23. For each of the four criterions, the differences between the L1 and L2 corpus are statistically significant (*Welch's t test* for (a): T = −3.487 (df = 23.000), $p < .01$; for (b): T = −2.438 (df = 25.240), $p < .05$; for (c): T = 4.593 (df = 3.933), $p < .05$; for (d): T = −3.405 (df = 23.000), $p < .01$).

In the following passage (11) produced by a L2 speaker, for example, the marked epistemic states of the speaker/narrator concerning the protagonist's inner world in clauses 05 and 06 indicate that the German learner of L2 Japanese keeps his own point of view, as is typical for native speakers of German:

(11)   gj10 (male, 24 years old): 05–07
   05   *sono-yatsu-wa totemo nodogakawaita-rashii-de*
      'The man seems to have become very thirsty, and'
   06   *tabun mizu-o sagashi-teiru-kara*
      '(he) is probably looking for water, for this reason'
   07   *ana-o tsukut-te*
      '(he) makes a hole, and […]'

4.4   Discussion

On grounds of the fact that the text-intrinsic shift in the temporal domain is the basis for temporal organization in L1 German as well as in L1 Japanese (see Section 3.3), it can be assumed that the L2 speakers of Japanese in the present analysis intended to organise anaphoric shift for establishment of temporal coherence. The analysis of the L2 corpus shows, however, that the L2 speakers do not consistently use the grammaticised concepts EVENT and STATE for this purpose. Instead, they rely on lexical means for temporal organization, just as in the retellings in their L1. This result suggests that L2 speakers have difficulties in discovering how the grammaticised concepts can be used and where they should be used to establish coherence, even though learners were taught the flexion and meaning of the aspect forms in the classes they took.

As the analysis of the L1 data showed, native speakers of German and Japanese follow different manners to involve the global topic entity in the conceptual structure for establishing coherence in texts (see points (c) and (d) in Section 3.3). Against this background, the result for the organization of the point of view by the L2 speakers suggests that the principles of information organization which are based on the notion POINT OF VIEW – in this case, the category that is grammaticised in the target language but not in the source language – are not evident to the learners mind and eye, and L2 speakers continue to rely on the principles of their L1. In other words: Not only is it hard to identify *how* and *where* to use the grammaticised notions (as opposed to more concrete notions expressed by lexical words), but implications of those notions for L1 speakers in establishing coherence are also difficult to discover.

Due to the scarce accessibility to the language-specific strategies for building up coherence in the target language, there might be two resources L2 speakers could

rely on during conceptualization processes namely, (a) the principles in their L1, and (b) concepts and logic which cannot be found in the source and target languages (see Bassetti & Cook 2011). The empirical data in the present study found supports for influences of the L1 principles (see Table 1). This, does not mean, however, that the possibility of (b) is precluded. In other words: It remains open whether L2 speakers who adopt the logic of coherence in their L1, may also develop their own 'new' strategies for information organization in order to reduce cognitive load for the highly complex task of retelling in L2.

5.   Final remarks

The text structure of narratives has been investigated since ancient times (see, for example the *hysteron proteron* 'latter before' in the traditional rhetoric), where the topic has preoccupied not only discourse analysts and rhetoricians but has also been examined from different interdisciplinary perspectives, such as anthropological or sociolinguistic standpoints (e.g., Grimes 1975; Longacre & Levinsohn 1978; Labov & Waletzky 1967; Dijk 1976; and numerous investigations within the framework of the functional sentence perspective in a broader sense). There are, however, only few studies which have looked at language-specific patterns of information organization at the text level by contrasting a corpus produced by speakers of typologically different languages (see the papers referred to in the introduction). This approach is, however, indispensable for an examination of the *language-on-cognition*-effect,[24] which has been one of the main focusses of psycholinguistic investigations in recent decades.

   The analysis in the present study, based on the typologically different languages German and Japanese, has shown that the grammaticised conceptual categories STATE, EVENT, and POINT OF VIEW influence conceptualization processes in L1 production, and further that the L1-related principles of information organization are persistent when speaking in L2. There are two extensions of the analysis framework which I would like to recommend for further investigation on the topic. The first extension concerns the typological features to be considered. Particularly, the feature [+/− subjectivity-prominent], which has hardly been addressed in the previous research framework, deserves further examination by using adequate language pairs. Furthermore, a fine-grained analysis of grammaticised notions would possibly suggest a re-typologization of languages, as was the case in the present study for the aspect languages English and Japanese.

---

24. *Cognition* here refers to cognitive processes engaged in preparing a 'preverbal message' to match particular communication goals.

The second extension concerns the type of content to be conveyed through language. For example, an empirical analysis of retellings of a story with three coequal protagonists showed that a shift in entity domain constitutes a basis for establishing coherence in Japanese texts, where a causal linkage of the type 'psychological cause – reaction' is extensively marked. In contrast, German speakers prefer the shift-in-time relation as the organizing principle, as was the case for the retellings of "Quest" (Tomita 2013).[25] Thus, speakers of Japanese seem to vary the principles for information organization regarding the number of animate entities involved in the story. It would be interesting to 'quest' for the integrative typological features that underlie the principles of information organization for different types of discourse in a language.

Acknowledgements

References

Bassetti, B. & Cook, V. 2011. Relating language and cognition. The second language user. In *Language and Bilingual Cognition*, V. Cook & B. Basetti (eds), 143–190. Hove: Psychology Press.

Berman, R. A. & Slobin, D. I. (eds). 1994. *Relating Events in Narrative: A Cross-linguistic Developmental Study*. Hillsdale NJ: Lawrence Erlbaum Associates.

Bohnemeyer, J. & Swift, M. 2004. Event realization and default aspect. *Linguistics and Philosophy* 27: 263–296.

Chafe, W. L. 1976. Giveness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and Topic*, C. N. Li (ed.), 25–56. New York NY: Academic Press.

Clancy, P. M. & Downing, P. 1987. The use of wa as a cohesion marker in Japanese oral narratives. In *Perspectives on Topicalization. The Case of Japanese wa* [Typological Studies in Language 14], J. Hinds, S. Iwasaki & Maynard S. K. (eds), 3–56. Amsterdam: John Benjamins.

Comrie, B. 1976. *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: CUP.

Dijk, T. A. van 1976. Philosophy of action and theory of narrative. *Poetics* 5: 287–338.

---

25. Tomita (2013) discusses these differences between the two speaker groups with respect to the typological opposition between [+/− topic prominent] (Li & Thompson 1976) in the following modified sense: German is a language with particles of information structure which when stressed highlight the polarity-value of a finite clause ("assertion-related particles", Dimroth et al. 2010; Klein 2012), whereas Japanese is a language with particles of information structure which mark the topic part of the propositional structure (i.e., the topic particle -*wa* and an additive particle -*mo*).

Dimroth, C., Andorno, C., Benazzo, S. & Verhagen, J. 2010. Given claims about new topics. How Romance and Germanic speakers link changed and maintained information in narrative discourse. *Journal of Pragmatics* 42: 3328–3344.

Fujinawa, Yasuhiro 2017. Licht und Schatten der kategorischen/thetischen Aussage. Kopula und Lokalisierungsverben im deutschen-japanischen Vergleich. In *Grammatische Funktionen aus Sicht der japanischen und deutschen Germanistik* [Linguistische Berichte Sonderheft 24], S. Tanaka, E. Leiss, W. Abraham & Fujinawa Y. (eds), 15–40. Buske: Hamburg

Grimes, J. E. 1975. *The Thread of discourse*. Paris: Mouton.

Ikegami, Y. 2005. Indices of a 'subjectivity-prominent' language: Between cognitive linguistics and linguistic typology. *Annual Review of Cognitive Linguistics* 3(1): 132–164.

Ikegami, Y. 2008. Subjective construal as a 'fashion of speaking' in Japanese. In *Current Trends in Contrastive Linguistics. Functional and Cognitive Perspective* [Studies in Functional and Structural Linguistics 60], G. González, M. de los Ángeles, J. L. Mackenzie & González E. M. (eds), 227–250. Amsterdam: John Benjamins.

Ikegami, Y. 2016. Subject-object contrast (shukaku-tairitsu) and subject-object merger (shukaku-gouitsu) in "thinking for speaking": A typology of the speaker's preferred stances of construal across languages and its implications for language teaching. In *Cognitive-Functional Approaches to the Study of Japanese as a Second Language*, K. Kabata & K. Toratani (eds), 301–318. Berlin: De Gruyter.

Jesperson, O. 1924. *The Philosophy of Grammar*. London: Henry Holt.

Klein, W. 1994. *Time in Language*. London: Routledge.

Klein, W. 2012. Assertion-related particles in German. In *Varietäten und Kontexte. Beiträge zu psycholinguistischen Schnittstellen*, K. Spalek & J. Domke (eds), 13–37. Tübingen: Stauffenburg.

Kudo, M. 1995. *Asupekuto-tensu taikei to tekusuto. Gendai nihongo no jikan no hyoogen* (Aspect-tense system and text: Temporal expressions in modern Japanese). Tokyo: Hitsuji.

Kuno, S. 1978. *Danwa no bunpoo* (Discourse grammar). Tookyoo: Taishuukan.

Labov, W. and Waletzky, J. 1967. Narrative analysis. Oral versions of personal experience. In *Essays on the Verbal and Visual Arts*, J. Helm (ed.), 12–44. Seattle WA: University of Washington.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: CUP.

Langacker, R. W. 1990. Subjectification. *Cognitive Linguistics* 1(1): 5–38.

Levelt, W. J. M. 1989. *Speaking. From Intention to Articulation*. Cambridge MA: The MIT Press.

Levelt, W. J. M. 1999. Producing spoken language: A blueprint of the speaker. In *The Neurocognition of Language*, C. M. Brown & P. Hagoort (eds), 83–122. Oxford: OUP.

Li, C. N. & Thompson, S. A. 1976. Subject and topic. A new typology of language. In *Subject and Topic*, C. N. Li (ed.), 457–489. New York, NY: Academic Press.

Longacre, R. E. & Levinsohn, S. 1978. Field analysis of discourse. In *Current Trends in Textlinguistics*, W. Dressler (ed.), 103–122. Berlin: De Gruyter.

Maynard, S. K. 1987. Thematization as a staging device in the Japanese narrative. In *Perspectives on Topicalization. The Case of Japanese wa* [Typological Studies in Language 14], J. Hinds, S. Iwasaki & Maynard S. K. (eds), 57–82. Amsterdam: John Benjamins.

Narita, T. 2009. *Shiten to nichi-dokugo no hyoogen* (Perspectiv and expressions in German and Japanese). *Tookyoo Gaikokugo Daigaku Ronshuu* 79: 399–414.

Okuda, Y. 1985a[1977]. *Asupekuto no kenkyuu o megutte. Kindaichi teki dankai* (On the study of aspect: The stage of Kindaichi). In *Kotoba no kenkyuu: Josetsu*, Y. Okuda (ed.), 85–104. Tokyo: Mugishoboo.

Okuda, Y. 1985b[1978]. *Asupekuto no kenkyuu o megutte* (On the study of aspect). In *Kotoba no kenkyuu: Josetsu*, Y. Okuda (ed.), 105–143. Tokyo: Mugishoboo.

Okugawa, I. 2007. *Katari no danwa ni okeru shiten to jitai ha'aku* (Viewpoint and cognitive construal in Japanese narrative discourse). *Tsukuba Ooyoo Gengogaku Kenkyuu* 14: 41–43.

Slobin, D. I. 1996. From "thought and language" to "thinking for speaking". In *Rethinking Linguistic Relativity*, J. J. Gumperz & S. C. Levinson (eds), 70–96. Cambridge: CUP.

von Stutterheim, C. 1997. *Einige Prinzipien des Textaufbaus. Empirische Untersuchungen zur Produktion mündlicher Texte*. Tübingen: Niemeyer.

von Stutterheim, C., Bouhaous, A., Carroll, M. & Sahonenko, N. 2012. Grammaticalised temporal categories, Language specificity and macroplanning in expository texts. *Linguistics* 50(2): 341–371.

von Stutterheim & M. Carroll, 2006. The impact of grammatical temporal categories on ultimate attainment in L2 leaning. In *Educating for Advanced Foreign Language Capacities*, H. Byrnes, H. Weger-Guntharp & K. Sprang (eds), 40–53. Washington DC: Georgetown University Press.

von Stutterheim, C., Carroll M. & Klein W. 2003. Two ways of construing complex temporal structures. In *Deictic Conceptualization of Space, Time and Person*, F. Lenz (ed.), 97–133. Amsterdam: John Benjamins.

von Stutterheim & Nüse, R. 2003. Processes of conceptualization in language production. Language-specific perspectives and event construal. *Linguistics* 41(5): 851–881.

Sugaya, N. & Shirai, Y. 2007. The acquisition of progeressive and resultative meanings of the imperfective aspect marker by L2 learners of Japanese. *Studies in Second Language Acquisition* 29: 1–38.

Teramura, H. 1984. *Nihongo no shintakusu to imi* (Japanese Syntax and Semantics), Vol. 2. Tokyo: Kuroshio.

Takemura, M. 2010. *Nihongo bogo washa to chuugokujin nihongo gakushuusha no danwa ni mirareru shiza* (The use of viewpoint expressions in the personal telling task and the picture description task by native Japanese speakers and Chinese learners of Japanese). *Hiroshima Daigaku Daigakuin Kyooiku-gaku Kenkyuuka Kiyoo* 2(59): 289–298.

Tomita, N. 2008. *Der Informationsaufbau in Erzählungen. Eine sprachvergleichende Untersuchung des Japanischen, des Deutschen und des Englischen zum Einfluss von einzelsprachlichen Systemeigenschaften auf die makrostrukturelle Planung*. München: Iudicium.

Tomita, N. 2013. Strategies for linking information by German and Japanese native speakers and by German learners of Japanese. *International Review of Applied Linguistics in Language Teaching* 51(2): 117–149.

Wei, Z. Z. 2010. *Jitai byoosha ni okeru taiwanjin nihongo gakushuusha to nihongo bogo washa no shiten no hikaku* (Comparison of the perspectives taken by Taiwanese learners of Japanese and native speakers of Japanese for description of events). *Gengo to Bunka. Nagoya Daigaku Daigakuin* 11: 255–270.

Texts as answers to questions: Information structure and its grammatical underpinnings in narratives and descriptions in German and English (topic and anaphoric linkage)

Christiane von Stutterheim and Mary Carroll
Heidelberg University

Text production requires the speaker to plan and formulate a coherent text in which information at the micro structural level is embedded and licensed by structures established at the macro structural level. The present paper addresses questions concerning the principles the speaker can rely on when solving a complex task of this kind. We assume that texts are organised around an initial-ising question, the *quaestio*, which introduces a space of alternatives which the speaker has to specify. The constraints set up by the *quaestio* leave the speaker with a range of options which in turn are constrained by language-specific principles partly rooted in the grammatical structure of the respective language. This claim is substantiated by a cross-linguistic study of English and German oral narratives. The analyses focus on grammatical features relating to the syntactic subject and word order constraints and aim to show how grammatical features which operate at the micro-structural level shape information organisation in macro-structural terms. The study also includes a comparison with object descriptions in English and German. They serve in illustrating the role of the underlying *quaestio* as well as grammatical categories, in this case the syntactic subject, which drives conceptualisation with transformations across the domains of space and entities when providing anaphoric linkage.

1. Introduction

Narratives are probably the most widely-studied type of texts. The approaches from which they have been analysed vary in many ways, both in terms of quantity as well as the range of theoretical perspectives from which they have been analysed. There is a basic definition which all approaches share: a narrative is an account of a chain of causally and temporally connected events. They are viewed as serving a

major function in the creation of identity, at a social as well as individual level, in that they make sense of the experienced world by transforming it into a coherent conceptual structure (McAdams et al. 2006; McAdams & McLean 2013). That is to say, narratives play an eminent role in making sense of the material and social world around us (see survey in Scheffel 2012; Herman et al. 2005). While narratives are viewed as a universal strategy of human cognition and communication, relatively little attention has been paid to differences at the level of the narrative strategy itself. In addressing this question it is useful to change perspective from product to process, from comprehension to production, and to ask what principles do speakers/writers follow in constructing a coherent text? In how far does the narrative task as such pre-structure content, coherence, as well as the use of linguistic means? Differences in narrative structure, as observed in cross linguistic studies, in particular in the context of first language acquisition, have been attributed to differences at the level of *rhetorical style* (Berman and Slobin 1994). However, the notion of rhetorical style implies a high degree of optionality, suggesting that a particular narrative form may be chosen on the basis of individual preferences, or particular traditions within one cultural community, as conveyed in institutions of learning, for example. Taking the grammatical structures of a language as the point of departure, cross-linguistic research strikingly reveals a systematic link between the language used and the way narratives are structured (Carroll et al. 2008). This factor is also evident in the study of second language acquisition, as speakers of a second language often fail to use the newly-acquired forms in accordance with the patterns of the target language (Carroll & Lambert 2003, 2006; Pavlenko 2005). These findings point to the fact that the respective linguistic system has to be considered as a crucial factor in structuring narratives in a specific way. In contrast to interpreting cross-linguistic differences as expressions of stylistic options, our analytical framework focuses on the impact of the respective linguistic system on narrative structure and in fact text structure in general.

Let us start with a short illustration comparing two translations from a passage of *War and Peace* by Leo Tolstoy (Book 5, Part 2, Chapter 16).

(1) *Угревшись   в своем угле, он заснул   и   перед* вечером
     Having got warm in his   corner he fell asleep and before evening
     только вышел   из балагана. *Nachdem er sich   in seiner Ecke*
     only   walked out of barack   After   he himself in his   corner
     ein bisschen erwärmt hatte, schlief   er sogleich ein,   und Ø
     a little   warmed had,   fell asleep he immediately and Ø
     kam erst   am Abend wieder aus   der Hütte <u>heraus.</u>
     <u>came</u> not-until at evening again   out of the hut   hither out.
     Having got warm in his corner he fell asleep and Ø <u>did not leave</u> the hut till toward evening.

At a first glance there is what may seem to be a marginal difference between the English and German translation with respect to the spatial perspective selected. In Russian and in the English translation the spatial perspective on the motion event '*Rostov go out of the hut*' is anchored at the protagonist, the subject of the sentence. In the German translation, the view on the situation is taken from the outside '**heraus**kommen'. The question is – why does the translator's decision vary in this way, although he could have found an equivalent at the level of linguistic form: 'hinausgehen/verlassen'? We find a first answer to this question when we consider the organisation of coherence in the different texts. The German text organises coherence via a shift in places 'what is happening inside', 'what is happening outside' while the Russian original and the English translation organise coherence around the protagonist. In the following we will demonstrate that information selection and information organisation is not just decided on at the level of the single sentence; they depend on predefinitions at the macro structural level. In other words, every narrative which provides a coherent account of an event sequence will be developed around a macro structural 'scaffold', so to speak. The central hypothesis underlying the present study is as follows: differences such as those found in the translations are, to a certain extent, due to language specific-constraints established at the level of macro structural organisation. Grammatical structures lead to over-arching constraints that hold on a default basis. In the following we will substantiate this hypothesis by comparing the role of grammatical constraints on information organisation in English and German texts.

2.   Texts as answers to questions

The idea of a macro structural level in information organisation is not new in linguistic studies of narrative texts. The first one to suggest this term was Bierwisch (1965) in his analysis of literary texts. In the seminal work on oral narratives by Labov and Waletzky (1967) we find the concept of a narrative *schema* which, in addition to the definition of what constitutes a narrative event, also encompasses different functional units: starting with the abstract, this is followed by the orientation, the complicating action, the evaluation, the result or resolution and finally the coda. While this schema presents the basis for differentiating different text types or genres, it did not set out to provide a framework for the integration of macro and micro structural planning within a global text structure. Van Dijk (1980) developed the idea of a textual macro structure further by formulating rules according to which texts can be generated, just as sentences can be constructed on the basis of grammatical rules. Since then the concept has been implemented in the field of text linguistics (Pearson & de Villiers 2009), as well as in research on different

types of cognitive linguistic impairments (Covington et al. 2005; Marini et al. 2008). All these studies showed how complex verbal productions of this kind require an integrated hierarchical form of organisation at a macro as well as micro structural level. However, there have been hardly any attempts to investigate language-specific patterns of macro structural organisation and the factors on which they are based. So the question of the specific effects of linguistic structure on macro structural planning does not arise.[1]

While most of the earlier work on macro structures in texts was carried out under the perspective of text comprehension, developments in psycholinguistics provided the basis for the analysis of information organisation in texts from the perspective of the speaker/writer. In production-oriented frames of analysis, global and local discourse structure is modeled as a multidimensional network, integrating syntactic, semantic, and pragmatic requirements. Central notions of text analysis are reinterpreted in this context under a dynamic perspective. Coherence is thus described in terms of information flow, with the integration of different cognitive processes involving the selection, segmentation, and linearisation of information stored in memory (Levelt 1989, 1996). But in order to understand coherence, we have to ask what is the integrating force which enables the speaker to plan and formulate a coherent text in which information at the micro structural level is

---

1.   The question of coherence at text level has also been taken up by formal semantic theories such as Discourse Representation Theory (Kamp & Reyle 1993), Rhetorical Structure Theory (Mann & Thompson 1988), Centering Theory (Grosz et al. 1995; Beaver 2002), or the QuD approach (*question under discussion*, Roberts 2012). Starting from the analysis of single sentences, formal theories were challenged by the fact that some linguistic phenomena could not be explained on the basis of the single sentence. Looking at phenomena such as anaphora resolution and tense/aspect structures, formalisms have to account for context dependency which can only be adequately described by assuming a structure beyond the single sentence. Formal approaches provide insights into specific selected phenomena of text structure. They give us highly elaborate and precise descriptions of the conditions under which anaphors are resolved. In Centering Theory, for instance, texts are viewed as dynamically changing states of attention: In order to explain what makes a text coherent, rules are formulated which structure the transition between sentences in terms of informational progression and salience of referents. The notion of macro structure is not used in this context. However, the rules which are formulated to describe the conditions under which texts can be evaluated with respect to their coherence properties operate at a global level. They underlie the decisions with respect to the informational status of the entities involved and the linguistic forms used. Thus we could say that CT provides a framework for formulating global constraints on discourse planning. But it also has its limitations. First, CT focuses on the domain of entities and their role in discourse coherence. Given the results of our cross linguistic studies this restriction seems to be rooted in a viewpoint based on coherence patterns that are typical for English. This relates to the second problem which lies in the assumption that rules for creating and understanding coherence will function across languages. Under this view language specificity is restricted to the level of linguistic representation and therefore it is not regarded as a relevant dimension in formal accounts of coherence.

embedded and 'licensed' by structures established at the macro structural level? What is the starting point in text production?

We assume that texts are organised around an initialising question which defines the task which the speaker has to solve. Thus, a narrative is the answer to a question such as: *what happened?* or, in more specific cases, *what happened yesterday?* A question of this kind introduces a space with alternatives from which the speaker will specify those which are assumed to hold as true, adequate and relevant in the light of the given question. Just as in the case of a single question-answer-pair, the over-arching question establishes constraints on what an adequate answer will entail. In the case of text production, the answer is distributed across a series of sentences or utterances, each one answering a sub-question which is derived from the central text question.[2] However, not all texts are initiated by an explicit question. The question which underlies the production of an utterance or a text may be explicitly asked, or it may be implicitly given – either by the speaker or by the nature of the situation as such. We have introduced the term *quaestio* for both cases – a concrete question in a real dialogue, or an implicit self-initiated question (cf. Klein & von Stutterheim 1987). This idea can be found in ancient rhetoric, although its structural consequences on the specific form of texts as well as individual utterances have rarely been a concrete object of investigation in the rhetorical tradition (cf. discussion of ancient rhetoric in von Stutterheim 1997).[3] The strong bias in earlier text-linguistic research on comprehension studies, which take the formulated text as the starting point of analysis, has obscured the view on the factors and principles which control the organisation of information. Most of the above-mentioned studies take the text as given and from this starting point set out to explain the form-function-relations as found in the sequences.

The *quaestio*-approach attempts to uncover some of the principles which guide the speaker in selecting and structuring information in order to reach the point of having a well-formed text.[4] While a single question-answer pair will show constraints that operate on a micro structural level, the macro structural level of information organisation comes into play if there are sequences of utterances with a

---

2. In the following we will use the term *utterance* to refer to the basic semantic unit which describes a situation. Since we are dealing with empirical oral data, the grammatical properties of a sentence might not always be fully maintained.

3. Quintilian defines the *quaestio* as the question at issue: *quaestio latius intelligitur omnis, de qua in utramque partem vel in plures dici credibiliter potest* and then elaborates on the text structuring function of the *quaestio* (1972: 390–392).

4. A related approach is the *Question Under Discussion* theory (Roberts 2012) which is also based on the idea of a text as an answer to a question. There is, however, a crucial difference to the *quaestio*-approach. The QUD theory assumes that language specificity is confined to the level of linguistic representation and that it has no implications for the generation of information structure.

global question these utterances are meant to answer. The structuring force of the *quaestio* thus concerns factors such as the information to be selected for expression, the choice of a relevant referential frame, patterns of coherence and linkage (which will be described as *referential movement*), information structure (topic/focus organisation). In other words, the *quaestio* introduces a macro structural frame which provides sufficient criteria for the speaker to construct information at the micro-level in accordance with requirements for coherence at a global level. In order to serve this function, the constraints put forward by the *quaestio* have to be specific enough in order to be implemented at the level of message planning. In this respect we distinguish between content- related and structural constraints. The latter encompass constraints on coherence as well as on information structure. Content-related constraints concern the type of information which has to be specified (*referential specification*), structural constraints concern the types of relations between sentences (*referential movement in the form of linkage*), as well as the distribution of topic-focus components at the macro- and micro structural level. This idea will be illustrated with an example of a *quaestio* which initiates a narrative text:

(2)   What happened at the battle of Borodino?

In answering this question, there are constraints which operate at the level of content in the selection and organisation of information. In describing these constraints, the information at issue is divided into five ontological domains:[5] *entity* and *predicate*, *time* and *space*, *modality*. Information at the sentence level can be described on the basis of these conceptual categories, whereby information components of each type can be recursively embedded. For the question above, constraints concerning the type of information which is relevant in each domain are formulated as follows:

The status of *modality* is factive, the domain *time* as well as *space* require information which is referentially located within a given temporal or spatial frame. For the domain of *entities*, the group of entities which are potentially relevant in selecting and structuring information is limited by the referential frame to a specific group – those entities which are involved in the battle of Borodino. In the *predicate* domain, event-type predicates are required as an answer to the question 'what happened?'. Sentences such as

(3)   a.   Napoleon inspected the battlefield in the morning.
     b.   Napoleon gave personal orders to his generals.

would be relevant and adequate answers at sentence level. We now come to the constraints that hold for the second dimension in information organisation, the

---

5.   Cf. a discussion on type and number of these domains von Stutterheim & Klein 2002.

question of linkage. We use the term *referential movement* in describing coherence relations across sentences in a text, given the nature of the information, and its development, within the respective referential domains. The following frame is established by the *quaestio* at this level: in the domain of modality the status is maintained, while in the temporal domain temporal references will be shifted, given the order of events, and thus form the domain in which the progression of information is defined. In the domain *space* constraints on referential movement are not strict, spatial references can be maintained or shifted or left unspecified. As regards the domain of *entities,* references can be introduced or maintained, while predicates have to introduce new information. Another important function of the *quaestio* lies in constraining information structure. As formulated above, the *quaestio* introduces a space of alternatives which have to be settled by the answer-text: This has been described in many different approaches, starting with H. Paul (1880) for the binary relation between question and answer.

(4)   Who fought in the battle of Borodino?

Those parts of the question which define the set of alternatives (in the present case the armies that could have fought in the battle of Borodino) form the *topic component* of the answer. The specification of members of the set of alternatives forms *the focus component*.

(5)   (The Russians and the French) focus (fought in the battle of Borodino) topic.

Topic and focus, as these terms are used here, are components of the 'meaning' of an utterance – *entities*, a *place*, a *time span*, an *action*. They do not refer to the level of linguistic form which expresses these meanings. When talking about the speaker's choice in the production process, we have to distinguish between the constraints on information structure set up by the *quaestio* and the selection of linguistic form to express these structures.[6] While this definition of topic and focus components is straightforward for the level of the sentence – and underlies the majority of studies on information structure (cf. Lambrecht 1994; Roberts 2012) – the nature of the constraints has to be cast in different terms if we are dealing with a text *quaestio*. In this case, the constraints will operate as topic/focus *conditions* at the macro structural level. These will then be implemented as local constraints on topic/focus components. In this view, the notions of discourse topic and sentence topic are defined on an identical basis. The introductory question sets up a space of

---

6.   The distinction between topic and focus must not be confused with the distinction between 'given' or 'maintained' information, on the one hand, and 'new' or 'introduced' information, on the other, although these dichotomies may often coincide (cf. the notion of *topic situation* in Klein 2008).

alternatives (cf. Rooth 1992) which in the case of a single sentence answer gives a topic-focus structure, with the elements defining the alternative space as the topic components and the specification functioning as the focus component. In the case of a text providing the answer to the question, the space of alternatives is defined by topic and focus conditions on a global basis, specifications which are then met at the level of the sentences by the specific elements which meet the conditions. Let us return to the example:

(6)   What happened at the battle of Borodino?

The question elicits a narrative text. It calls for the specification of some complex event – we will call this the macro event – which the speaker has to subdivide into a series of sub-events, each located (a) within some time interval ti, which is part of the time interval of the macro event, (b) within a macro space (s_m), (c) within the real world (w_{fac}). Thus, the *quaestio* for the entire text can be subdivided into a set of sub-*quaestiones* for which the following topic conditions hold:

1:   What happened to whom at t1 within s_m and in $w_{fac?}$
2:   What happened to whom at t2 within s_m and in w_{fac}?
3:   What happened to whom at t3 within s_m and in w_{fac}?
n: What happened to whom at tn within s_m and in w_{fac}?

For each individual utterance, we have two types of topic condition: (a) those which involve the maintenance of properties of the alternative defining components. These are termed static components of coherence; (b) those which concern conditions of change are termed dynamic components of coherence, as, in the case of narratives, with temporal shift in the domain of *time*. This forms the basis for progression and provides the 'backbone' of the narrative. Taking the example above, we could, for example, insert an utterance (7b) which does not interrupt coherence as such, but it does not meet the macro structural constraints set up by the *quaestio* for the expression of temporal shift.

(7)   a.   Napoleon inspected the battlefield in the morning.
      b.   He considered this to be absolutely indispensable.
      c.   Later he gave personal orders to his generals.

The second sentence leaves the time line, and although information of this kind can be highly relevant, it does not meet the condition for temporal shift, as well as the predicate type constraint, and therefore does not answer the question *what happened at t*n. Utterances which directly contribute to answering the question are defined as constituting the *main structure* of a text, and those which give additional – communicatively often important – material its *side structure*. But note, what counts as the *main structure*, or *side structure*, is relative, and depends on the specific *quaestio*.

While we have taken narratives for purposes of exemplification, the constraints outlined above apply by analogy to other text types such as descriptive or argumentative texts. These constraints define the constitutive macro structural properties which function as a scaffold for implementing information at the micro structural level, i.e. the level at which information and the respective expressive devices are selected for any question-answer constellation:

(a) the categories specified in the relevant conceptual domains (e.g. *time*, *space*, *entity*)
(b) referential movement within the domains from one main structure utterance to the next, including both static and dynamic components
(c) the assignment of specific meaning elements to the topic component or the focus component
(a)–(c) the main structure of a text; information selected which does not comply with these constraints functions as a side structure, and has to be specifically integrated

These constraints narrow down a specific range of options for the speaker – but they also leave options open with ways to 'present the case'. A major factor in this context is given with language-specific constraints on how the speaker may proceed in selecting information to meet the criteria specified. In the following we will show that linguistic structure has an effect on information structure, leading to language-specific patterns.

3.   Language specificity in information organisation

Constraints which are constitutive for a specific text type leave room for options at the macro structural level with regard to both the selection of information as well as information organisation.

In what follows, we will show how language-specific constraints on information organisation are partly rooted in grammaticalised categories. The cross-linguistic analyses carried out to substantiate this claim involve narratives in German and English. The analyses focus on one grammatical component which differs for the two languages: the syntactic subject, taking into account its formal features concerning word order constraints, as well as semantic features. The study includes a comparison with object descriptions in English. This serves to illustrate the role of the underlying *quaestio* and the way the domain of *space* is profiled in information organisation in narratives in contrast to object descriptions.

3.1   Macro structural planning in narratives and object descriptions in English and German: the role of grammar

The role of planning procedures at a macro level for information organisation will be illustrated in the following on the basis of the domains of *space* and *time*. We will show how grammatical properties of a language, in this case syntactic constraints on the constituents that can occur in first position in the sentence, affect information organisation. English and German differ in the linguistic means with ready access to the first position in the clause (cf. von Stutterheim & Carroll 2005). These grammatical constraints, however, interact with principles of information organisation, leading to convergence or divergence depending on whether the domains under focus, *space* and *time*, have a global topic status or not.

Object descriptions differ from narratives in that the domain of *space* is profiled for information progression by the global topic condition given with the *quaestio*. This is the case when describing an object (e.g. the layout of a town), with the requirement of specifying the spatial relations which hold between individual entities (e.g. *where is what?*). This status does not apply to the domain of *space* in narratives, although descriptions of the setting in which events occur often require specification of spatial relations between the entities involved in the events. In a narrative the domain of *time* is one topic component in information organisation (*what happened at t1, t2 …?*), and topic status applies to temporal relations between events and the temporal sequence in which they occur, as outlined above. The contrast with respect to the role of the domain of *space* in both contexts is at the focus of the present section: spatial relations play a function in providing anaphoric linkage in both narratives and descriptions. Their status differs, however, depending on the *quaestio* and language-specific grammatical constraints on constituents that typically occur in clause-initial position, the privileged position in information organisation in establishing anaphoric relations and linkage across sentences in English (Fox 1993; Lambrecht 1994; Beaver 2004; Klein 2008). This position brings into focus the language-specific implementations of the global constraints on information organisation as set up by the relevant *quaestio*. The present window on information organisation is thus given by grammatical differences between English and German and how the languages differ in the linguistic means with ready access to the first position in the clause (cf. Carroll et al. 2008; von Stutterheim & Carroll 2005, 2007; Carroll et al. 2000; Carroll & von Stutterheim 2011). The relevant factor for English is the relatively fixed position of the syntactic subject, with the subject as the main constituent for mention in the first position preceding the verb. This is not the case in German as the position preceding the finite verb can be taken up by any major constituent:

(8) a. *Asterix und Getafix sind gestern   in Londinium angekommen.*
          Asterix and Getafix are   yesterday in Londinium arrived.
     b. *Gestern   sind Asterix und Getafix in Londinium angekommen.*
          Yesterday are   Asterix und Getafix in Londinium arrived.
     c. *In Londinium sind gestern   Asterix und Getafix angekommen.*
          In Londinium are   yesterday Asterix and Getafix arrived.

All three patterns are grammatical in German, while (8b) and (8c) are ungrammatical or restricted in English. In the V2 language German, the 'verb-second constraint' (V2) confines the first position preceding the verb to a single main constituent, but does not constrain the type of constituent that can occur in this position, as opposed to English. Significantly, the first constituent preceding the verb need not be the syntactic subject, which can precede or follow the verb, in contrast to English. These grammatical differences between German and English are reflected in the status accorded to spatial information in the role of anaphoric linkage when describing locations in the context of a narrative, as opposed to an object description. As will be shown in the next section, in English the syntactic subject remains the dominant constituent in providing linkage in descriptive passages which have *side structure* status in narratives; occurrences of adverbials or prepositional phrases encoding spatial information are not accorded access to clause-initial position, with clear consequences for information organisation. In object descriptions by contrast, spatial relations have access to clause-initial position. In each case we will show how the constraints observed hold for the task as a whole, drawing on empirical data to illustrate the theoretical claims.

3.2   Descriptive passages in film re-narrations in English and German

Native speakers of English and German (20 speakers per group) were asked to retell the content of a short silent film (*Quest*, approx. 7 minutes long). It portrays the adventures of a clay figure, the only animate-like protagonist, and his encounters with a series of natural forces (high winds, flying rocks, huge sheets of paper) on his quest for survival in a hostile environment. The participants were told that they will first see the entire film, which will then be re-played and stopped after each main episode (5 in all). Their task was to tell *what happened?* in the episode they just saw. The data base consists of 20 English retellings with an average length of 173 utterance units and German retellings with an average length of 185 utterances.

3.2.1   *English*
Starting with a feature which both languages share, spatial information on the places and circumstances in which the events occur can be encoded in *clause-final* position, as in 'the clay figure lands *in a new world*'. This is observed in both the

English and German narratives when the protagonist and the event mentioned are located for the first time. The relevant differences occur in linkage to the subsequent sentence. The main cross-linguistic difference in this context is as follows: when spatial information could serve the function of providing linkage to the preceding utterance there is a switch to a noun phrase in this role in English in the narratives. In this case the constituent that ensures linkage is the syntactic subject.

(9)   English
      you see him land <u>on a new plane</u>
      this plane is full of paper

Linkage by means of a pronoun is illustrated in the following examples.

(10)   the man has fallen <u>onto another bleak landscape</u>
       it looks as if it is all made out of paper

(11)   so now the man has dropped from the sky <u>onto the next land</u>
       this land has a desert aspect
       it just stretches to the horizon
       but this time the floor is made of paper
       it's a very windy land

With linkage via a noun or pronoun there is a switch in information structure across domains from the domain of space, as given with the locative prepositional phrase, to that of entities, as expressed by the syntactic subject. Significantly, shifts of this kind are not observed in German.

3.2.2   *German*
The linguistic means used in similar contexts in the German data are typically spatial relations encoded in prepositional phrases or adverbs, with linkage in this form occurring in clause-initial position (Example (12), (13)).

   Linkage via prepositional phrases

(12)   German
       und das Männchen schlägt <u>auf einer</u> <u>Oberfläche</u> <u>auf</u> und *auf dieser*
       and the little man bangs <u>on a</u> <u>surface</u> <u>on</u> and on this
       Oberfläche herrscht sehr großer Wind
       surface reigns very strong wind

Linkage via adverbials, *da* (there), *überall* (everywhere)

(13)   aus dem Sandloch raus fällt er <u>in so ne Blätterwelt</u>
       out of the sandhole out falls he into a kind of 'sheets of paper' world
       und *da* hört er dann wieder das Wasser tropfen
       and there hears he then again the water dripping

(14) jetzt sieht man eine Ebene *überall* fliegen   auch Papierblätter
      now sees   one   a   plane everywhere fly   also   sheets of paper
      durch   die Luft
      through the air

The cross-linguistic differences point to the way in which linkage is managed. In contrast to German, information in the domain of *space* is not accorded access to sentence-initial position in English, as shown by the predominance of the syntactic subject in this role. The restriction can thus be attributed to the syntactic subject as the main candidate for mention as the first constituent in the clause, as well as the status accorded to *space* in information structure in a narrative, which does not have global topic status. If we take a look at the topic domain of *time,* however, the patterns show that constraints observed for spatial relations do not apply to information concerning temporal relations between events.

3.3   Temporal relations in film re-narrations in English and German

Temporal relations, as expressed by forms such as *then*, *later,* have access to sentence-initial position in English as well as German, based on the status accorded to temporal relations with the *quaestio* (*what happened at t1, t2 …?*). Temporal relations constitute the central topic condition in information organisation for this text.

(15) but this time the floor is made of paper
      it's a very windy land
      and some sheets of paper blowing around
      there's a little tornado of paper
      and he gets up
      and dusts himself off again
      gets hit in the face by a flying paper
      he's knocked onto the floor
      and then he suddenly gets up …

(16) German
      und er fängt an dort auch nach Wasser zu suchen streckt   die
      and he begins   there also for   water   to search   stretches the
      Hände gegen   `n Himmel als wenn er schauen wollte   oder
      hands toward   the sky   as   if   he to see   wanted or
      fühlen wollte   ob Regen kommt *und dann* fängt   er eben auch
      to feel wanted if   rain   comes and then begins he also
      wieder   an dieser nassen stelle zu graben
      again   at this   wet   place to dig

To summarise the observations so far, access to the first position in the clause is not observed for spatial information in the narrative texts in English. A constraint of this kind does not hold in German and spatial adverbials or prepositional phrases occur in sentence-initial position to provide linkage. We assume that there are two factors that account for the different patterns in information structure in the English and German narratives (i) the status ascribed to information concerning the event sequence in a narrative, given the underlying *quaestio,* and (ii) constraints on word order. With regard to status, *temporal relations* between events meet the global topic condition for information progression. This holds in both languages so that temporal information can occur in the first position of the sentence, thereby preceding the syntactic subject in English and the finite verb in German. The first cross-linguistic difference is that access to the position preceding the subject in English is not accorded to information specifying *spatial relations* between entities, as outlined above. In English, access to sentence-initial position is observed almost exclusively for the syntactic subject in this context. Significantly for the role of macro planning principles, the status ascribed to *spatial relations* between entities differs in tasks involving an object description in English. In contrast to narratives, they constitute a core referential domain which is subject to topic conditions for information progression. As will be shown below, it takes 'topic status' to propel information involving spatial relations between entities into the position preceding the subject in English, as observed in object descriptions given the quaestio *where is what*?

As mentioned above with regard to German, the 'verb-second constraint' confines the first position preceding the verb to a single main constituent. In English, by contrast, constituents have to be accommodated, so to speak, as the subject also precedes the verb (post-positions are rare). In German this situation is solved grammatically: although topic status profiles one option over the other for mention in clause-initial position, other constituents can be placed in sentence-initial position, given the structural feature whereby access is confined to only one constituent. This promotes the basis for linkage via clause initial position at a *local level* - a paramount feature of verb-second languages such as German with regard to information structure (see conclusions below). Spatial relations have access to clause initial position in a narrative, but this occurs only if related information is also mentioned in the preceding utterance, as in Example 13 above. This restriction does not hold for spatial relations between entities in object descriptions: they have global topic status in this context, on the basis of the underlying quaestio, with access in overall terms to clause initial position as the core category in providing linkage (see Example 25 below).

3.4   Object descriptions

In this task, speakers of English and German (14 speakers of English; 15 speakers of German) were asked to describe a picture of an old town. They were asked "to describe the town with the layout of the area so that the interlocutor would have an idea of where the buildings and other objects mentioned were located relative to one another". The interlocutor was present, but did not sit facing the picture, which was on the wall in front of the speaker. It is necessary to ask speakers specifically to locate the objects they describe, as this is not always a core feature of a description. In the narrative task, by contrast, there is no need to ask the speakers specifically to supply the temporal relations between events, as the order of events is given with the established nature of the task. As they were specifically asked to locate the objects they identify in their description, the underlying *quaestio* is to tell *where is what*? Apart from two cases, there was no insistence on maintaining this *quaestio*. This procedure was carried out in order to determine the extent to which speakers of English assign topic status to spatial information or not. In other words, the option of assigning topic status to the domain of entities by describing *what is where*? was not ruled out (with the exception of the two speakers who were explicitly instructed). In the case of *what is where*, which profiles the domain of entities, the syntactic subject is the main candidate for mention as the first constituent in the clause and not spatial information.

In contrast to descriptive passages in narratives, spatial information has access to clause-initial position on a systematic basis in the English as well as the German object descriptions. Nine of the fourteen speakers of English organise information on the basis of *where is what*? Of the five who proceeded on the basis of *what is where*? the interlocutor intervened in two cases by asking the speaker to supply information on locations on a consistent basis (see below); the underlying *quaestio* then changed to *where is what*, as evidenced in the data. Both patterns occur consistently in tasks of this kind in English (see in detail Carroll, Murcia-Serra, Bendiscioli & Watorek 2000).

–   With '*what* is where' (five speakers), the description is presented as concerning a macro-entity (e.g. *this picture looks to me like a section of an old town*) which is broken down into sub-entities. The domain of *entities* has topic status and information progresses as changes in attention in this domain (there is an entity x, there is an entity y etc). The first position in the clause is typically occupied by the syntactic subject. Progression via spatial relations is mainly confined to adjacency relations between entities at a local level (see Example 21 below, 008 *next to x is y*).

–   In the case of '*where is what*' (nine speakers), the feature at the centre of attention is the picture as a global 'space' (*starting in the middle*). In this case the domain of *space* is accorded topic status and information progresses as changes in attention in the spatial domain. Progression is not based mainly on adjacency relations between entities but can also progress across any space on the picture (e.g. from the foreground to the background of the picture).

(17)   quaestio *where is what?*
    001   ok *starting on the right side* there are/ there's a number of houses
    002   forming a U-shape (….)
    010   the block starts out really narrow
    011   it looks like it expands *down along the stream along the river*
    012   *next to it over the stream* is a bridge
    013   *in front of the building and over the stream* there are some people and a guy on a bicycle
    014   *in the background where the stream kind of heads toward* is some more buildings
    015   *and along the whole horizon* is a mountain, along the entire picture, almost the entire picture

The topic condition with *where is what*? is met by linking information via a shift in the spatial domain which is then typically placed in first position. With the space given by the entire picture as the topic domain, progression is marked by frequent use of word orders with a verb-second pattern (see Example 17, 012, 014, 015). A further example of a V2 pattern is shown in the following description.

(18)   quaestio *where is what*? (verb second word order)
    001   in the middle of the picture is a square
    002   that has a fountain and some trees
    003   *running in front of the square* is a street car and a truck and some cars
    004   also *in the square* is a fruit stand
    005   and *in front of the fruit stand* is a little kiosk

This word order is highly restricted in general in English, but its occurrence in the present data is apparently given with the status ascribed to spatial information over the domain of entities (see discussion below).

The intervention asking speakers to change their *quaestio* was carried out on a preliminary basis with two of the five speakers who consistently proceeded on the basis of *what is where*? The emphasis was thus placed on entities and their features and less on providing information on their location. Following the intervention, where the speakers were asked to tell *where is what*?, the word order pattern for spatial information changed from predominantly clause-final, following mention of the entity, to clause -initial position.

(19) *what is where*? (before intervention)
    001 okay this picture looks to me like a section of an old part of a European city
    002 judging from the architecture
    003 it's a street scene (..)
    004 starting with the building on the left
    005 it looks like a house (…)
    025 in the next part of the picture there's what looks like a square
    026 there's a statue
    027 it's a public area
    028 there are trees (..)
    032 there's a kiosk for buying tobacco and magazines

Intervention at this point in the description: *Could you say a little more like where things are in relation to one another, that would be helpful?*

(20) following intervention
    001 okay you mean at the square, just where they are in relation to each other?
    002 okay there is a kiosk
    003 where you can buy magazines
    004 *to the left of that* there are three children
    005 the girl *right next to the kiosk* is wearing green
    006 *to the left of her the girl* is wearing pink
    007 and *to the left of her* there's a girl, looks like she is wearing blue
    008 *to the right of the kiosk* there is an old man with a beard
    009 looks like he is feeding the birds
    010 which are in front of the kiosk
    011 *behind the kiosk* there's another market

In asking the speaker to tell *where is what*? spatial information is placed in clause-initial position in the descriptions. Although this test is preliminary, it confirms the predominant order observed in the other descriptions classified as providing an answer to *where is what*?

Examples of typical texts based on *what is where*? (five speakers) run as follows:

(21) *quaestio what is where*?
    001 okay *this picture* looks to me like a section of an old part of a European city
    002 *its* a street scene
    003 *it* has several different streets running perpendicular to the front of it
    003 starting with *the building* on the left
    004 *it* looks like a house
    005 there's a business *on the bottom floor* (…)
    008 *next to that in front of that building* there is a man using a jackhammer
    009 there's someone loitering *in the doorway of the building*
    010 *next to that* there's a fresh air market with people selling and buying fresh produce

The predominant pattern in information structure is to place references to entities in clause initial position (nouns, pronouns). Spatial information occurs in only two of the main clauses in clause-initial position (Example 21, 008, 010). A similar pattern applies in the following Example (007):

(22) 001 I see a group of different types of buildings *along two different streets*
       002 *it* appears to be a quiet street
       003 although there are several cars and a street car and a truck
       004 but *they* don't appear to be going very fast
       005 there are a lot of different tiny little shops,
       006 first a grocery shop a candy shop a tobacco shop and a restaurant
       007 *in the distance* apparently there are residential areas
       008 because you can see people working *at their window*
       009 and the children playing *in the playground*
       010 there's a quiet street *off the main street*

The descriptions are dominated by the introduction of entities via presentationals (*there is..*, it is …) as well as nouns and pronouns as subject of the clause in clause initial position. It should be noted that in English *there is* is not a locative but an empty subject (e.g. *there is a bus stop over there*). Spatial information is more likely to occur in sentence-final position if the place is maintained ((21) *the house* in 005, *the building* in 009). Given a shift in place ((21) *next to that* in 008 and 010), spatial information can occur in sentence-initial position in providing anaphoric linkage.

In the descriptions in German thirteen out of fifteen follow *where is what?* and not *what is where?*

(23) German
       001 also jetzt sind wir in einer Stadt
              now are we in a   town
       002 gehen wir mal von links vor
           proceed we   from left
       003 da   sind Häuser ziemlich verwinkelte Dächer
           there (loc) are houses quite   elaborated roofs
       004 auf der Terrasse über Steiners Weinhandlung wahrscheinlich
           on the terrace   over steiners wine shop   probably
           spielen Kinder   Blinde Kuh
           play   children blind man's buff
       005 und dahinter   ist auch ein alter Mann mit
           and there-behind is also an old   man   with
           einem Stock
           a   walking stick

In the two descriptions that do not specifically provide an answer to *where is what?* linkage is mainly provided dynamically by *dann kommt x* (then comes x); there is a focus on entities with mention in clause-final position (e.g. *dann kommt ein Fachwerkhaus*, then comes a timber-framed house).[7]

(24)   001   also es ist offensichtlich in einer Stadt
                   it   is   obviously   in a   town

        002   man steht also vor   einer Straße
                one   stands   before a   street

        003   und sieht die Straße also praktisch quer
                and   sees   the street   also practically sideways

        004   und dann zunächst einmal die Häuser
                and   then   at first   the houses

        005   die   an dieser Straße stehen
                which at this   street   stand

        006   da   ist ein Haus mit einem Restaurant unten drin (…)
                there (loc) is a   house with a   restaurant below inside

        009   dann kommt ein kleiner Bach
                then   comes   a   small   stream

        010   dann ein Fachwerkhaus
                then   a   timber-framed house

        011   dann kommt eine schmale Gasse
                then   comes   a   narrow   laneway

The following excerpt presents an example with spatial information in sentence initial position. Occurrences in this context are not confined to the presence of relation information in the preceding clause (009), given the underlying *quaestio*, in contrast to narratives.

---

7.   Sections focussing entities by means of *es gibt* (it exists) can also be observed in the descriptions in German but occurrences are low (e.g. es gibt eine Straßenbahn, einen kleinen Platz mit einem Brunnen und einen kleinen Kiosk wo man Tabak kaufen kann; ziemlich nah an der Hauptstraße befindet sich ein kleines Café; *there is a tram, a small square with a fountain and a small kiosk where one can buy tobacco; pretty close to the main street finds itself a small cafe*).

(25)   006   man   kann   einige   kleine   hinterhöfe   erkennen   mit
one   can   some   small   courtyards   recognise   with
kleinen   gartenanlagen
small   gardens

007   direkt   in   der   mitte   zieht   sich   ein   kleiner   fluß   durch
directly   in   the   centre   stretches   itself   a   small   river   through
dieses/diesen   teil   der   altstadt
this   part   of   the   old-town

008   der   von   einer   brücke   gekreuzt   wird
which   from   a   bridge   crossed   is

009   **überall**   sind   leute   auf   den   straßen   unterwegs
everywhere *are   people   on   the   streets   out and about*

In sum, the German and English descriptions differ with regard to the frequency with which speakers adopt the sub-questions *where is what?* versus *what is where?* when organizing information for expression in an object description, although both groups were asked to tell *where is what*? The occurrence of *what is where*? in English in five out of fourteen descriptions may be traced to the grammatical differences between English, an SVO language, in which the syntactic subject is the principal constituent for expression in sentence-initial position, and German, a V2 language, in which this is not the case. Nevertheless, *where is what*? is the dominant pattern in English in the object descriptions, with frequent use of a V2-like word order when the domain of *space,* given with the entire picture, is the topic domain for information progression. Interestingly, the use of presentationals, a form which explicitly introduces and focuses an entity (*there is an x*) is low in the descriptions in which the domain of *space* has topic status (*in the square is a fruit stand*). The status accorded by the quaestio '*where is what*' to the domain of *space,* rather than *entities,* may thus explain the absence of '*there is an x*', with its entity-based focus, in the V2-like orders (*over the stream is a bridge* rather than *over the stream there is a bridge*). It should be mentioned that use of V2 is highly restricted in present day English to certain predicate types, in particular in dynamic contexts, and does not correspond to V2 status in German. However, the frequency of occurrence of this variant in English in the present contexts reveals its role in information organisation.

4.   Discussion and conclusions

The comparison between narratives and object descriptions provides empirical evidence for two factors which serve as the integrating force for text production - factors which are hierarchically ordered. At the higher level there is the initialising *quaestio* which accords topic status to the relevant ontological domains. At a subordinate level there is the linguistic system with grammatical constraints which are language specific. The empirical analysis provides evidence for the way in which patterns in information organisation at the micro level are licensed at the macro structural level in texts, but also proceed in conjunction with language specific constraints.

In the case of an object description, the *quaestio* to which the speaker will provide an answer can be formulated as *where is what*? for example. The ontological domain of *space,* and the specification of spatial relations between entities, has topic status and determines information progression. Significantly for the present study, this status does not apply for the expression of spatial relations in a narrative context. With regard to the *quaestio* in a narrative context, speakers will provide an answer to the question *what happened to x at tn?* This means the ontological domain of *time* is profiled at a macro structural level, as speakers will provide an answer involving temporal relations between events. Spatial relations between entities are not irrelevant, but they do not constitute the core ontological domain in which information progresses, which in a narrative is *time* and not *space.* The global constraint established at the level of macro structural planning leads to different patterns in the status accorded to spatial information in the narratives and descriptions in English and German, given grammatical differences in word order. In the narratives in the SVO language English, access to the initial position in providing linkage holds for the syntactic subject of the clause in passages where temporal shift is not relevant as a linkage strategy, but this does not apply to spatial information. Information organisation proceeds as follows in this context: Instead of the specification of explicit spatial relations between entities, as observed in similar contexts in the German narratives, descriptive information on spatial components of the situation is encoded in nominal form. Although speakers could, in theory, present information on a location as follows, given

(26)   entities, e.g. rocks (x) as located at plane (y): *on this plane* (y) *there are huge rocks* (x),

locatives do not function in providing anaphoric linkage. There is a transformation to the domain of entities with maintenance of the noun, in the form of the syntactic subject, thus giving a place of the type x:

(27)   *you see him land on a new plane* (y) *this plane* (x) *is full of huge rocks* (x).

In other words, areas in space, so to speak, are thus represented as entities (*this plane*) and not as locations (*in this place, on this plane*).

The relevance of macro structural planning in information organisation is exemplified by the fact that this constraint does not apply in English to spatial relations in the context of an object description. In this case, prepositional phrases or adverbials encoding spatial relations have access to the first position in the clause in anaphoric function. In other words, information progresses via spatial relations on a systematic scale in this context. It does not have to be nominalised in order to access sentence initial position and provide linkage. In this case, patterns of information organisation converge in English and German.

Concerning the first position in a main clause, all languages show some degree of competition of categories for the first position (Lambrecht 1994). The function assigned to the subject in English in linkage in narratives is given by the fact that although English may allow adverbials to precede the subject in order to link information, their use in this context is restricted. In its development from a V2-like structure in Old English to its present form, the subject has been endowed with the scope of encoding content that can link existing propositions to the previous set, along with the expansion of other means such as passives, it-clefts (see Traugott 1972; Denison 1985). However, we see that in order to meet the requirements in information organisation for the expression of spatial relations in an object description, speakers of English implement a verb-second pattern, with the syntactic subject following the verb, thus ensuring access to clause-initial position for means that express spatial relations. The speakers draw on a sequence which is more typical of its roots as a verb-second language. Although overt linking of one utterance to the next is not typical of English, compared to verb-second languages (Virtanen 1992), occurrences of the V2-like pattern are observed in the object descriptions presented above. This pattern is no longer frequent in present-day English, but we see how frequency of use increases in accordance with task demands at the level of information structure.

The comparison also underlines how English will frequently draw on the domain of *entities* in ensuring discourse coherence, given the status of the syntactic subject in a language with a relatively fixed word order. Although anaphoric linkage via the subject is a core feature in establishing linkage for English, as profiled in Centering Theory (Beaver 2004), it should be emphasised this does not hold for languages with a relatively free word order, as in the V2 language German, for example. Spatial and temporal relations have access to sentence-initial position and occur freely in narratives and descriptions in ensuring linkage. The domain of *entities* plays a secondary role given the grammatical options in German.

The overall consistency with which context-dependent patterns are observed in text production in the cross-linguistic comparison underlines the presence of language-specific guidelines at a global level. Moreover, decisions on constituents and their eligibility for mention in specific positions in the clause need not be verified anew for each utterance as it arises. Macro structural plans are available to assist with information organisation and information load, allowing the speaker to proceed on a default basis where required. From a cognitive-linguistic point of view, information organisation involves language use in the production of what may be called 'large scale coherence' (cf. Jackendoff 2002; von Stutterheim & Klein 1989). Given its multidimensional complexity, we need to refine and extend research on long-standing propositions concerning 'knowledge at the level of 'conceptualiser', and 'formulator' in the context of language production models (Levelt 1989), or knowledge at 'interfaces' between the different sub-modules of a language (Jackendoff 2002). In the present study, for example, we gain a glimpse of how syntactic categories, in this case the syntactic subject, drive conceptualisation, with transformations across the domains of *space* and *entities*, for example, in order to carry out the task of anaphoric linkage.

Although different methodological approaches shed light on different aspects of research on narrative structure, the underlying linguistic knowledge concerning overarching principles of information organisation has to be taken into account even when analyses of anaphoric relations are mainly confined to adjacency relations over a limited number of clauses. The concept of an initialising question or *quaestio* provides a tool to bring this knowledge to the surface. The *quaestio* functions as an integrating force that enables the speaker to plan and formulate a coherent text in which information at the micro structural level is embedded and 'licensed' by structures established at a macro structural level in text planning.

Acknowledgements

References

Beaver, D. 2002. Presupposition projection in DRT: A critical assessment, In D. Beaver, L. Casilas Martinez, B. Clark & S. Kaufmann (eds), *The Construction of Meaning*, Stanford: CSLI Publications.

Beaver, D. 2004. The optimization of discourse anaphora. *Linguistics and Philosophy* 27: 3–56.

Berman, R. A. & Slobin, D. I. (eds). 1994. *Relating Events in Narratives: A Cross-linguistic Developmental Study*. Hillsdale NJ: Lawrence Erlbaum Associates.

Bierwisch, M. 1965. Poetik und Linguistik. In *Mathematik und Dichtung*, H. Kreuzer & R. Gunzenhäuser (eds), 46–66. München: Nymphenburger.

Carroll, M., Murcia-Serra, J., Bendiscioli, A. & Watorek, M. 2000. The relevance of information organisation to second language acquisition studies. *Studies in Second Language Acquisition* 22(3): 441–466.

Carroll, M. & Lambert, M. 2003. Information structure in narratives and the role of grammaticised knowledge: A study of adult French and German learners of English. In *Information Structure and the Dynamics of Language Acquisition* [Studies in Bilingualism 26], C. Dimroth & M. Starren (eds), 267–287. Amsterdam: John Benjamins.

Carroll, M. & Lambert, M. 2006. Reorganizing principles of information structure in advanced L2s: A study of French and German learners of English. In *Educating for Advanced Foreign Language Capacities*, H. Byrnes, H. Weger-Guntharp & K. Sprang (eds), 54–73. Washington DC: Georgetown University Press.

Carroll, M., Roßdeutscher, A., Lambert, M. & von Stutterheim, C. 2008. Subordination in narratives and macrostructural planning. A comparative point of view. In *'Subordination' versus 'Coordination' in Sentence and Text. A Cross-linguistic Perspective* [Studies in Language Companion Series 98], C. Fabricius-Hansen, H. Weger-Guntharp & W. Ramm (eds) ,161–186. Amsterdam: John Benjamins.

Carroll, M. & von Stutterheim, C. 2011. Event representation, time event relations, and clause structure: A crosslinguistic study of English and German. In *Event Representation*, E. Pederson & J. Bohnemeyer (eds), 68–83. Cambridge: CUP.

Covington, M. A., He., C., Brown, C., Naci, L., McClain, J. T. & Fjordbak, B. S. 2005. Schizophrenia and the structure of language: The linguist's view. *Schizophrenia Research* 77: 85–98.

Denison, D. 1985. The origins of periphrastic DO: Ellegård and Visser reconsidered. In *Papers from the 4th International Conference on English Historical Linguistics: Amsterdam, 10–13 April 1985* [Current Issues in Linguistic Theory 41], R. Eaton, O. Fischer, W. Koopman & F. van der Leek (eds), 45–60. Amsterdam: John Benjamins.

Dijk, T. van 1980. *Macrostructures. An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition*. Hillsdale NJ: Lawrence Erlbaum Associates.

Fox, B. 1993. *Discourse Structure and Anaphora*. Cambridge: CUP.

Grosz, Barbara J., Joshi, Aravind K. & Weinstein, S. 1995. Centering: A framework formodeling the local coherence of discourse. *Computational Linguistics* 21: 203–226

Herman, D., Jahn, M. & Ryan, M. -L. (eds). 2005. *Routledge Encyclopedia of Narrative Theory*. London: Routledge.

Jackendoff, R. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: OUP.

Kamp, H. & Reyle, U. 1993. *From Discourse to Logic*. Dordrecht: Kluwer.

Klein, W. 2008. The topic situation. In *Empirische Forschung und Theoriebildung: Beiträge aus Soziolinguistik, Gesprochene-Sprache und Zweitspracherwerbsforschung: Festschrift für Norbert Dittmar*, D. B. Ahrenholz, U. Bredel, W. Klein, M. Rost-Roth & R. Skiba (eds), 287–305. Frankfurt: Peter Lang.

Klein, W. & von Stutterheim, C. 1987. Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte* 109: 163–183.

Labov, W. & Waletzki, J. 1967. Narrative analysis. In *Essays on the Verbal and Visual Arts*, J. Helm (ed.), 12–44. Seattle WA: University of Washington Press.

Lambrecht, K. 1994. *Information Structure and Sentence Form*. Cambridge: CUP.

Levelt, W. J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge MA: The MIT Press.

Levelt, W. J. M. 1996. Perspective taking and ellipsis in spatial descriptions. In *Language and Space*, P. Bloom, M. A. Peterson, M. F. Garrett & L. Nadel (eds), 77–107. Cambridge MA: The MIT Press.

Mann, W. C. & Thompson, S. A. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8(3): 243–281.

Marini, A., Spoletini, I., Rubino, I. A., Ciuffa, M., Bria, P. & Martinotti, G. 2008. The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research* 105: 144–155.

McAdams, D. P., Josselson, R. & Lieblich, A. 2006. *Identity and Story: Creating Self in Narrative*. Washington DC: American Psychological Association.

McAdams, D. P. & McLean, K. C. 2013. Narrative identity. *Current Directions in Psychological Science* 22(3): 233–238.

Paul, H. 1880: *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.

Pavlenko, A. 2005. Bilingualism and thought. In *Handbook of Bilingualism: Psycholinguistic Approaches*, A. B. M. de Groot & J. Kroll (eds), 433–453. Oxford: OUP.

Pearson, B. & de Villiers, P. 2009. Discourse, narrative and pragmatic development. In *Concise Encyclopedia of Pragmatics*, K. Brown (ed.), 206–216. Oxford: Elsevier.

Quintilian, M. F. 1972/1975. Ausbildung des Redners, 12 Bücher, H. Rahn (ed.). Darmstadt: Wissenschaftliche Buchgesellschaft.

Roberts, C. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5(6): 1–69.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1: 75–116.

Scheffel, M. 2012. Nach dem narrative turn. Handbücher und Lexika des 21. Jahrhunderts. *Diegesis* 1(1): 43–55.

von Stutterheim, C. 1997. *Einige Prinzipien des Textaufbaus*. Tübingen: Niemeyer.

von Stutterheim, C. & Carroll, M. 2005. Subjektwahl und Topikkontinuität im Deutschen und Englischen. *Zeitschrift für Literaturwissenschaft und Linguistik* 139: 7–27.

von Stutterheim, C. & Carroll, M. 2007. Durch die Grammatik fokussiert. *Zeitschrift für Literaturwissenschaft und Linguistik* 145: 35–60.

von Stutterheim, C. & Klein, W. 1989. Referential movement in descriptive and narrative discourse. In *Language Processing in Social Context*, R. Dietrich & C. F. Graumann (eds), 39–76. Amsterdam: Elsevier.

von Stutterheim, C. & Klein, W. 2002. Quaestio and L-perspectivation. In *Perspective and Perspectivation in Discourse* [Human Cognitive Processing 9], C. F. Graumann & W. Kallmeyer (eds), 59–88. Amsterdam: John Benjamins.

Tolstoi, L. 1868 *Krieg und Frieden*, translated by M. Kegel 2010. München: Winkler.

Tolstoi, L. 1868. *War and Peace*, translated by C. Garnett 2004. New York: Random House.

Толстой, Л. 1996. Война и мир. Собрание сочинений в восьми томах Т. 3,4. Москва: "Лексика".

Traugott, E. 1972. *A History of English Syntax*. New York NY: Holt, Rinehart, and Winston.

Virtanen, T. 1992. Given and new information in adverbials: Clause-initial adverbials of time and space. *Journal of Pragmatics* 14: 99–115.

Discourse prominence and the selection of anaphora – evidence from pronouns in historical German

Svetlana Petrova

Bergische Universität Wuppertal

Based on the analysis of the distributional properties of different types of pronouns in historical German, the present paper investigates discourse prominence as a factor governing the selection of anaphora in the discourse. It claims that particular types of pronouns refer to entities which are more salient than others not only with respect to the immediately preceding context, but also with respect to the global prominence of the respective referents in the entire discourse, including the subsequent context.

1. The role of discourse prominence in the choice of anaphora – introductory remarks

The choice between different classes of expressions allowing resumption of referents[1] already present in the discourse is a fundamental component of natural languages. There is general consensus on the issue that the selection of one particular type of expression out of a set of alternatives depends on the activation status, also called the salience of the referent in the mental state of the interlocutors (Ariel 2001; Almor & Nair 2007; Arnold 2010, among others). Salience, in turn, becomes particularly important if the context provides more than one potential antecedent to resolve the anaphor, and depends on a number of various, interrelated factors. Some of them affect the semantic properties of the referent, such as animacy, thematic role or lexically-driven preferences arising in different expressions of interpersonal relations (Ariel 1988; Cozijn et al. 2011). Others have been linked to the role of the referent in the earlier discourse, such as the grammatical role or

---

1. I basically focus on anaphora resuming referents of nominal expressions. But of course, whole propositions may act as antecedents as well (Asher 1993).

the information-structural value of the antecedent in the previous utterance (cf. Section 2 for discussion and references).

Recent work on anaphora has added the referent's role in the global discourse, including the subsequent context, to the factors affecting the salience of the referent (see Arnold 2010). Garnham and Cowles (2008) accurately mirror this issue by creating the impressive image of the Janus-like behavior of anaphora in the discourse, whose basic property is to "look" both backwards, to the contents of the previous utterance, but also forwards, to the way in which the discourse is intended to proceed.

This perspective on the behavior of anaphora attaches to the notion of *discourse prominence*, also called *discourse potential* of referents, which Givón (1983) introduces to indicate the referents' importance in the global context, including the following one. Discourse prominence is a correlate of the referent's persistence, measured in terms of frequency of explicit mention of this referent in a sequence of clauses after introduction. In the current linguistic discussion, discourse prominence has proved to be a significant factor in the explanation of a variety of phenomena of linguistic representation, such as the competition between canonical and non-canonical types of indefinite expressions[2] or the use of special, marked word order types.[3] According to the literature, the common property underlying the contribution of the special forms or patterns is that they highlight a referent as a major protagonist and subject of subsequent mention, in contrast to the canonical types of phrases or constructions which are indifferent with respect to the discourse potential of the referents that they involve.

The present paper will employ the notion of discourse prominence in order to examine the referents' discourse potential as a factor in the selection of anaphora. The evidence on which the investigation is based upon comes from historical German, which displays an inventory of anaphoric pronouns which goes beyond

---

2. Take the use of special types of indefinites across languages, e.g. the phonologically reduced cognates of the numeral 'one' in Russian (Ionin 2013) or in spoken Hebrew (Givón 1981), the indefinite use of *this* in English (Wright & Givón 1987; Ionin 2006) and German, including the novel colloquial German form *so'n* (Deichsel 2011; Deichsel & von Heusinger 2011).

3. Discourse prominence in terms of expected and intended topichood has been associated with syntactic phenomena, e.g. with various marked word order patterns or special syntactic constructions. In German, which is a canonical OV language, extraposition of arguments and adjuncts has been explained as a strategy of singling out the constituent that provides the designated topic of the following clause (Vinckel 2006; Vinckel-Roisin 2011). Exactly the same effect has been detected as one of three basic conditions licensing multiple XP-fronting in German root clauses (Bildhauer & Cook 2010). In this case, the designated discourse topic rests in its base position in the clause-internal domain but the remaining constituents are evacuated to the pre-finite field, causing putative violations of the well-known verb-second constraint for German root clauses.

that of modern German (Behaghel 1923: Vol. I, 275–348). On the basis of corpus data from Old High German (OHG), the paper will examine the referential properties of different classes of pronouns, i.e. of personal pronouns vs. demonstratives on the one hand, and of different classes of demonstratives on the other. By comparing the frequency of resumption of entities introduced by one of these different competing means of anaphoric reference, the paper aims at providing insights into the role of discourse prominence in the selection of pronouns in discourse. This, in turn, bears implications for our understanding of the role of anaphoric reference in narrative discourse. The paper will claim that the forward-looking function of pronouns is a means of enhancing information processing by determining the salience relations among different potential topics in the subsequent discourse.

2. Previous ways of determining salience in pronominal choice and resolution

As already outlined in Section 1, reference to antecedents can be established by means of different lexical types of expressions, such as definite DPs, various classes of pronouns, e.g. demonstratives, personal pronouns, and zero elements. It is obvious that these types of anaphora differ regarding their degree of lexical explicitness, see (1), with definite descriptions being highly explicit, zero pronouns being least explicit, and demonstratives and 3rd person pronouns being in the middle of this scale (see also Almor & Nair 2007: 86):

(1)   definite DP – demonstratives – personal pronouns – zero

It is commonly assumed that there is an inverse relation between the degree of lexical explicitness of an anaphor and the referent's activation status, or salience in the mental state of the interlocutors. Broadly speaking, the less salient a referent, the more explicit the anaphor that is used to resume it. As to what determines the activation state, or the salience of the referent, various properties of the antecedent have been investigated in the literature, such as the grammatical function of the antecedent in the previous clause, the semantic nature of the referent or its information-structural value in the context. It has also been observed that some of the properties determining the salience status of a referent only hold as long as the referring expression is used within one and the same "thematic paragraph" (Givón 1983: 7). This is because episode boundaries, e.g. the beginnings of new paragraphs or sections, are taken to minimize the salience status of referents, regardless of the properties of the antecedent in the immediately preceding sentence (see Ariel 2001; Grüning & Kibrik 2005).

One of the earliest accounts on salience, couched in the framework of the centering theory and applied to the modeling of anaphoric relations in discourse, is the one developed by Brennan et al. (1987). The main claim that this approach puts forward is that the salience of a referent correlates with the grammatical function of its antecedent expression in the argument structure of the governing verb in the immediately preceding utterance. More precisely, this framework suggests the salience scale represented in (2). According to it, the most salient referent is the one whose antecedent expression is realized as the subject of the immediately preceding sentence. It is followed by the referent of the internal object, i.e. the XP that the lexical verb V° takes as its sister. Less salient are the referents of higher arguments, e.g. indirect objects selected by the intermediate projection involving the verb and its internal object ([V° + XP]), as well as those realized as adjuncts and modifiers:

(2)   subject > object [1] > object [2] > adjuncts > others

The centering approach to salience was challenged on the basis of counterevidence suggesting that the grammatical function of referents may be overridden by other factors, more specifically by ones pertaining to information structure. On closer examination of the principles governing pronominal reference in German and English, Strube and Hahn (1999) propose to replace grammatical function by *functional information structure*, i.e. the informational status of the referent in the preceding discourse. They apply a triple distinction of informational statuses, invoking parallels to the taxonomy of given-new put forward in Prince (1981):[4]

(3)   hearer-old > mediated > hearer-new

Finally, Bosch and Umbach (2007) present data showing that both grammatical function and givenness can be overridden by a third factor, which in their interpretation pertains to topicality. In the discourse in (4), the demonstrative may resume only one of two equally pre-mentioned, thus hearer-old referents, despite the grammatical function of their antecedents in the previous clause:

(4)   *Woher Karl das weiß?*
      'How does Karl know about that?'
   a.   *Peter hat es ihm gesagt. Der{Peter/*Karl} war gerade hier.*
        'Peter has told it to him. DEM was here a while ago.'
   b.   *Er hat es von Peter gehört. Der{Peter/*Karl} war gerade hier.*
        'He has heard it from Peter. DEM was here a while ago.'

---

4.   Note that on a broader empirical basis, going beyond pronominal reference, Gundel et al. (1993) have worked out a more extended givenness hierarchy, which distinguishes six cognitive statuses correlating with different lexical forms of referring expressions, including indefinite and definite nominal phrases and different types of pronouns.

Bosch and Umbach (2007) assign the following interpretation to these facts. Although the first conjunct in (4a–b) provides two hearer-old referents (*Karl* and *Peter*) as potential antecedents of the *d*-pronoun, only one of them is already pre-established as the discourse topic of the respective context. This is the crucial difference regarding the information-structural properties of the two referents, and the decisive factor in determining the salience relation between them. In general, Bosch and Umbach (2007) claim that the most salient referent is the already established discourse topic in the respective context. As a demonstrative is lexically more explicit than a personal pronoun, it resumes the less salient referent, namely the non-topical referent *Peter*, and rejects the highly salient one, the already established discourse topic *Karl*.

Although the above mentioned approaches differ in the way in which they define salience, they are similar in that they base their analyses on properties of the antecedents in the previous discourse. Garnham et al. (1996) and Garnham and Cowles (2008) shift the attention towards the role of the subsequent context in the interpretation of anaphora. They show that in contexts involving verbs of interpersonal relations, the reference of the pronoun is resolved only after processing the contents of the following utterance, while the context preceding the pronoun might be completely identical. Note that the pronoun *he* is understood as resuming two different referents in (5a) and (5b) respectively, only depending on our knowledge of the causality effects implied by the verbs in the second conjuncts:

(5) a. Bill confessed to John because **he**{Bill/*John} wanted a reduced sentence.
     b. Bill confessed to John because **he**{*Bill/John} offered a reduced sentence.

The question arises if the following context is also relevant beyond cases in which world knowledge determines the resolution. More precisely, we want to know if the intended prominence of a referent in the following context plays a role in the selection of the type of referring expression used to resume it.

If this hypothesis is supported by empirical evidence, then the choice between different lexical classes of referring expressions will constitute another domain in which, next to variation regarding the semantic type and the syntactic realization of nominal expressions, discourse prominence will come into play. In the next section, I will report on the results of two case studies carried out to examine the role of the discourse potential of referents regarding the choice of anaphora in data from historical German.

3.   Case studies

3.1   Data and method

According to Behaghel (1923: Vol. I, 280–348), anaphoric relations in OHG can be established by means of one of the following lexical types of pronouns.[5,6]

(6)   a.   3rd person pronouns *(h)er, si(u), iz,* modern German *er, sie, es*
       b.   the simple demonstratives *ther, thiu, thaz,* modern German *der, die, das*
       c.   the compound demonstratives *theser, thisiu, this,* modern German *dieser, diese, dieses*
       d.   the distal demonstrative *(g)ener, (g)eniu, (g)enez,* modern German *jener, jene, jenes*
       e.   the identity markers *sama* 'the same one' (only relics of it are attested in OHG) and *selb* 'the same one', modern German *derselbe* 'the afore mentioned person/object'[7]
       f.   the demonstrative *sulig* 'such, of this kind', modern German *(ein) solcher*

Behaghel (ibid.) provides some suggestions regarding the distribution and the referential properties of these different types of pronouns but in general, the variation displayed by the data has remained unexplored in later research. Only recently, the properties of single representatives have been investigated in some detail. Demske (2005) analyzes the changes in the lexical properties and the consequent shift in the inflectional behavior of the demonstrative *sulig* in the history of German. Additionally, the variation between 3rd person pronouns and demonstratives in OHG has attracted some attention, see Solf (2008), Petrova & Solf (2010) and Speyer (in this volume). On the complex interaction of different types of pronouns in Middle High German (*c.* 1050–1350, henceforth MHG), see Haferland (2013).

Two types of contexts allow access to exploring the discourse behavior and the referential properties of anaphora in OHG. First, we can examine the contexts displaying competition between different types of anaphora used to resume a single entity present in the context, as e.g. the use of 3rd person pronouns and *d*-pronouns exemplified in (7) und (8):

---

5.   I only provide the nominative singular forms of these pronouns, but of course, I also take into consideration the plural forms as well as the respective forms in the oblique cases. The full paradigms of the individual pronouns, including doublets, are given in Braune (2004), *inter alia.*

6.   I neglect empty pronouns in this investigation. On the status of OHG as a potential *pro*-drop language, see Axel (2007) and Axel & Weiß (2011).

7.   The meaning of identity with another object is a later development.

(7) *inti arstuont sliumo **tház magatin**/ Inti gieng: **siu**   uúas alt/*
and rose up   quickly the girl   and walked PERS.PR was old
zuelif iaro.
twelve years

'And the girl stood up and walked. She was twelve years of age.'

Lat. & surrexit continuo puella/ & ambulabat: erat autem/annorum duodecim.
(T 96, 30–32)[8]

(8) *uuas thô thâr **anna uuizzaga**/ dohter   fanueles   fon*
was then there Anna prophetess daughter Phanuel-GEN from
*cunne   aseres,/   **thiu** gigienc fram In managa taga.'*
tribe-DAT Aser-GEN DEM went   forth in many   days

'There lived the prophetess Anna, a daughter of Phanuel of the tribe of Aser, and she was of a great age.'

Lat. & erat anna proph&issa/filia fanuel de tribu aser,/hæc processerat in diebus multis.   (T 38, 22–24)

But in addition, we can take contexts in which different types of anaphora are needed to resume different individuals simultaneously present in the discourse. E.g., in (9), the personal pronoun *er* and the demonstrative *therer* refer to two different individuals activated in the discourse, namely to Adam and the Son of God:

(9) *Thiz íst min **sún diurer** [1][…] / **Ádam**[2] er[2]   firkos*
this is my son dear   Adam,   PERS.PR betrayed
*mih […] / ih wane, **thérer**[1] fulle   állaz thaz ih wílle.*
me   I believe DEM   fulfill-PRES.SUBJ all   that I want

'This is my dear son […] Adam, he betrayed me […], I believe that this one [the Son of God] will do all I want.'   (O I, 25, 17–20)

In two subsequent case studies, I will consider both types of contexts to investigate the conditions that license anaphora in OHG. In particular, I will be interested in the role of the referents' discourse prominence in the choice of the anaphor. In Case Study I, I will test if there is a difference in the discourse prominence of referents resumed by different pronouns in those cases in which reference to single referents is established. I will take as a subject of investigation the variation between personal pronouns and simple demonstratives and compare the discourse prominence of referents which are resumed by each of these pronouns within 10 subsequent clauses.

---

8. The slash '/' in the examples taken from the OHG translation of the *Tatian* diathessaron represents a line break. '&' substitutes 'et'.

In Case Study II, I will test if the difference in the discourse prominence of two referents simultaneously present in the context correlates with the type of pronoun that is selected to distinguish them. Here I will concentrate on instances in which the distal demonstrative *jener* is combined with another type of anaphor, e.g. with a personal pronoun or a demonstrative of any kind.

I collected a corpus of examples from OHG by using the word searching tool of the TITUS[9] database. For the purposes of Case Study I, I extracted examples in which members of the formal paradigm of the simple demonstrative pronoun of the type in (6b) occur in the OHG *Tatian* translation.[10] Since this pronoun can also occur in adnominal position, functioning as a forerunner of the modern German definite determiner, and because it also acts as a relative pronoun in attributive relative clauses, I scanned the results of the electronic search and manually selected those instances in which the respective pronoun occurs independently, as an anaphoric expression. A total of 40 instances were considered for the study. In a next step, I collected the same number of 3rd person pronouns appearing in the same text under the same conditions.

For the purpose of Case Study II, I searched the OHG texts in the TITUS database for all instances of the pronoun *jener* in independent anaphoric use and selected those in which this pronoun occurs in combination with another anaphoric expression, e.g. a personal pronoun or a demonstrative of any kind. Examples that meet these criteria are not very frequent in the data. I found three relevant examples in Otfrid's *Gospel Book* and eight in Notker's works. Admittedly, this data is very sparse, but its analysis is worthwhile because, as will be outlined in the discussion of the examples, the interpretation of the cases attested in OHG reveals a coherent picture of the discourse properties of different types of pronouns with respect to the discourse prominence of the respective referents.

In both case studies, I implemented the method of measuring referential persistence established in Givón (1983) and applied in Wright & Givón (1987), Deichsel (2011), Deichsel & von Heusinger (2011) and related work. Basically, this method consists in determining the number of explicit mentions of a referent in a series of

---

9.   Thesaurus Indogermanischer Text- und Sprachmaterialien, see http://titus.fkidg1.uni-frankfurt.de.

10. In this case study, I confine myself to the *Tatian* translation because I intend to draw on some of my own previous observations on the distribution of 3rd person pronouns and demonstratives in this text (see Petrova & Solf 2010). Otfrid's poetic work, another major text of the OHG attestation, was left aside in Petrova & Solf (2010) and will be also neglected here because it displays its very own, specific distribution of the same types of pronouns which at this point cannot be explained along the lines of what is seen in *Tatian*. See also Speyer (in this volume), who concentrates on Otfrid's poetic work exclusively.

subsequent clauses.[11] The number of clauses considered in the previous literature varies among authors. I checked 10 clauses after the anaphoric resumption by the respective anaphor.[12] Note that I also counted as explicit mentions resumptions by means of a null pronoun, if this element fills an argument position in the argument structure of the governing verb.

3.2   Case study I: 3rd person pronouns vs. simple demonstratives in OHG

In OHG, 3rd person pronouns and simple demonstratives are similarly used to resume an antecedent in the previous discourse. This is illustrated in (10) and (11):

(10) *thin   quena elysab&h / gibirit thir   **sun**. / Inti nemnis   thû*
      your   wife   Elizabeth bear   you-DAT son   and name-2SG you
      *sinan   namon   Iohannem. / Inti **her**   ist thir   gifeho*
      his-ACC name-ACC John-ACC   and PERS.PR is   you-DAT joy
      Inti blidida
      and gladness
      'Your wife Elizabeth will bear you a son, and you will call his name John, and he will be joy and gladness for you.'
      Lat. & uxor tua elysab&h/pari& tibi filium./& uocabis nomen eius Iohannem/& erit tibi gaudium & exultatio.   (T 26, 25–28)

---

11. Additionally, topic persistence, i.e. the mention of the respective referent as the topic of subsequent utterances, is measured in these studies, but I will neglect this value because of the various theoretic and conceptual problems related with the notion and the determination of sentence topics.

12. Note that in Case Study I, I selected only those examples in which one of the respective anaphora resumes the referent for the first time after introduction by a lexical DP. This means that I neglected instances like the demonstrative *then*-DAT.PL in examples like (i) because here, this is not the first anaphoric mention of the referent after introduction. Rather, the scribes and the Pharisees, the antecedent of the demonstrative, have been resumed by a personal pronoun (*in*-DAT.PL) in the previous sentence:

(i) *gisahun tho **thie buohhara inti pharisei** / thaz her áz mit den*
     saw   the the   scribes   and   Pharisees that He   ate with the
     *suntigon/[…] ther heiland quad **ín**/[…]   **then** quad hér*
     sinful-DAT.PL the   Saviour said PERS.PR DEM said he
     'The scribes and Pharisees saw that He ate with the sinners […] The Saviour told them […] To them He said.'
     Lat. & uidentes scribę & pharisęi/quia manducar& cum peccatoribus/& publicanis dicebant/[…] hoc audito ihesus ait illis […] quibus ipse ait'   (T 91, 13–31)

(11) *Seno   nu   Inphahis   In reue   Inti gibiris   **sun**. / Inti*
behold now conceive-2SG in womb-DAT and bear-2SG son   and
*ginemnis sinan   namon   heilant, / **ther** ist mihhil.*
name-2SG his-ACC name-ACC Heiland DEM is great
'Behold, you shall conceive in your womb and give birth to a son and shall call
his name Jesus. He shall be great.'
Lat. ecce concipies In utero & paries filium./& uocabis nomen eius ihesum./
his erit magnus.   (T 28, 17–19)

From the perspective of minimal pairs like (10) and (11) above, the distribution of
the two types of anaphora looks accidental, because the properties of the anteced-
ents, including their semantic nature and their lexical environment, are completely
identical. Note also that the antecedents in (10) and (11) behave equally with respect
to each individual factor discussed in the literature, i.e. both antecedents function
as non-subjects in the argument structure of the same verb, they share the same
thematic role and the same activation state, and fail to act as an already established
discourse topic in the respective context. Additionally, as Solf (2008) and Petrova
and Solf (2010) observe, there is no strict correspondence between OHG demon-
stratives and special classes of pronouns in the system of Latin, e.g. *hic, haec, hoc*
or *is, ea, id*. In other words, OHG displays the same variation as the one observed
in modern German and illustrated by virtue of examples like (12):

(12) ***Eine Frau** kam herein. **Sie/Die** trug einen roten Mantel.*
'A woman came in. PERS.PR/DEM wore a red coat.'

Descriptive grammars, e.g. the 5th edition of Duden (1995, 333), explain the choice
of the demonstrative as a typical property of colloquial usage, attributing the entire
variation to factors pertaining to register and style. But a series of recent inves-
tigations (Strube & Hahn 1999; Abraham 2002; Bosch et al. 2003, 2007; Bosch
& Umbach 2007) has shown that such an explanation fails to account for some
well-known interpretational effects that occur if the context provides more than
one potential antecedent, like in (13):

(13) ***Paul** wollte mit **Peter** laufen gehen, aber **er**{Paul/Peter}/**der**{*Paul, Peter}war
krank.*
'Paul wanted to jog with Peter, but PERS.PR/DEM was ill.'
(Bosch et al. 2007: 146)

The personal pronoun can be interpreted as resuming each of the two referents
mentioned in the preceding sentence, i.e. the one of the subject expression (Paul)
and the one of the PP (Peter). By contrast, the demonstrative is more limited in its

reference, in that it only refers to the referent of the PP (Peter) but strongly rejects reference to the referent of the subject expression (Paul). In other words, there is a truth-conditional effect related to the use of the two types of anaphora that is related to the fact that the demonstrative eliminates ambiguities that may arise if a personal pronoun occurs in contexts which provide more than one potential antecedent.

Bosch and Umbach (2007) have shown that the grammatical function of the antecedent is not the decisive property that makes demonstratives reference infelicitous. As already discussed in Section 2, they show that demonstratives may resume subject antecedents as well, provided that they are not established as discourse topic in the respective context.

Solf (2008) and Petrova and Solf (2010) investigate the role of salience in the selection and interpretation of 3rd person pronouns and demonstratives in OHG. They consider a complex of factors including the role of the antecedent expression in the previous context, foremost its grammatical role, its informational status and its topichood. The result of the analysis shows that the selection of anaphora in OHG proceeds systematically, in a regular fashion. Personal pronouns, which are less explicit, take up referents which, compared to other referents present in the context, have the highest salience score regarding the factors considered. E.g. in (14), the personal pronoun *he* resumes the already established discourse topic, the blind man:

(14)   Jesus heals **a blind born man**[1] who used to sit and beg in the street. **A group of people**[2] seeing this man starts a discussion wondering if he is the same person they used to know. **Some of these people**[3] claim:
'He is it [= the same blind born beggar]'
Her[1]   ist íz.
he   is it
Lat. hic est   (T 221, 5)

By contrast, demonstratives take up referents which are less salient, compared to other referents present in the discourse. Consider (15). Two referents are present, namely a dead man and his mother. Both referents are mentioned for the first time in the immediately preceding sentence, and none of them represents an already established discourse topic. But they differ with respect to the grammatical function of their antecedent in the previous clause. While the dead man is realized as the subject of the clause, the mother is introduced by a phrase functioning as a genitive modifier. Thus, the latter is less salient than the former one. In this case, the demonstrative resumes the less salient one of the two referents:

(15) *senu   **arstorbaner[1]** / uúas gitragan einag sun / sinero   **muoter[2]***
Behold dead man   was   carried   single son   his-GEN mother
*Inti **thiu[2]** uuas uuituuua*
and DEM   was   widow
'Behold, a dead man was being carried out, the only son of his mother, and
she was a widow.'
Lat. ecce defunctus/efferebatur. filius unicus/matris suæ. & hæc uidua erat.
(T 84, 22–24)

The distributional principles exemplified for contexts involving a single anaphoric
expression are maintained if we take examples with two different referents resumed
by personal pronouns and demonstratives. Consider (16):

(16) **a master**[1] calls **one of his slaves**[2] and demands a certain explanation, **the
latter**[2] begins to speak
ther[2]** tho   quad **imo[1]
DEM   then said   him-DAT.SG
'He told him.'
Lat. isque dixit illi.   (T 156, 18)

The first referent (the master) is highly salient because it is introduced earlier and
provides the already established discourse topic. Additionally, its antecedent ex-
pression is the subject of the previous clause. The second referent (the slave) is
less salient because it is mentioned for the first time in the immediately preceding
sentence and is therefore non-topical. In addition, its antecedent expression is the
object of the governing verb in the previous clause. Thus, the personal pronoun
imo-DAT.SG resumes the more salient referent (the master), while the demonstrative
ther resumes the less salient one (the slave).

There are contexts in which the choice of the pronoun is crucial for the reso-
lution of the anaphor and for the truth-conditional interpretation of the utterance.
Solf (2008) and Petrova and Solf (2010) discuss the example in (17) in which there
are two potential plural antecedents for the demonstrative *dea* 'them': (i) *sine angila*
'his angels', which is given and realized as the subject of the previous sentence, and
(ii) *alle dea (a)suuihi enti dea ubiltatun* 'all the sins and the evil-doers', which is novel
and takes the function of the object in the previous clause. Clearly, only the second
referent can provide the proper antecedent of the demonstrative: not the angels but
the evil-doers will be cast into the fire. Consequently, the demonstrative is used
because it unambiguously refers to the less salient of the two potential antecedents:

(17) **The Son of Man**[1] will send out **His angels**[2] and **they**[2] will take **all the
sins and those who do evil things**[3]
*enti tuoit   **dea** {*[2]/[3]} in fyures   ouan*
and do-3SG.PRES DEM-ACC.PL in fire-GEN.SG oven
'and [the Son of Man] will send them into the fire.'   (MF X, 5)

Let us add to this picture the role of the referent in the following context and examine the question if the referent's discourse prominence is a factor governing the choice between personal pronouns and demonstratives in OHG. In order to study this relation, I compared the discourse persistence of referents which are resumed by personal pronouns or by demonstratives for the first time after introduction by a full DP. I counted the number of explicit mentions of these referents by anaphoric means within 10 subsequent clauses after they have been resumed by an anaphor.

Figure 1 shows how often referents resumed by a 3rd person pronoun (left-hand column) and by a demonstrative (right-hand column) are referred to in the following discourse. The numbers of explicit mentions, starting from no mention (0x) and ending with seven or more than seven mentions (> 7x) are given in the bottom line.

Referential persistence of referents of 3rd person pronouns vs. demonstratives in OHG
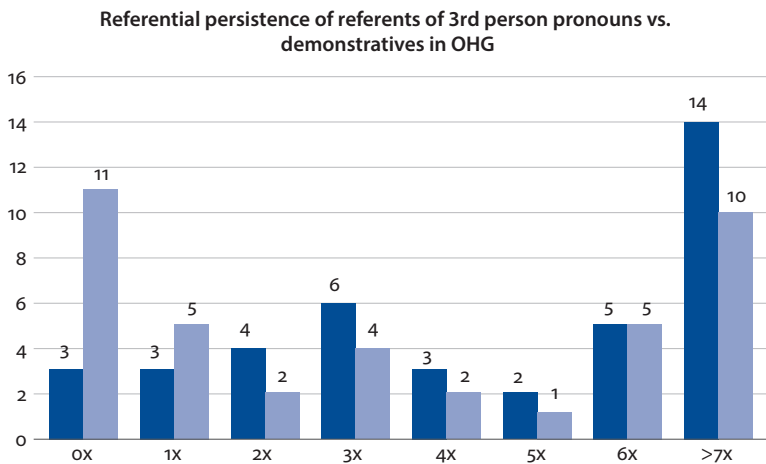
Figure 1. Number of resumption of referents of 3rd person pronouns and demonstratives in 10 subsequent clauses (total values, 40 clauses containing each anaphor)

The two endpoles in Figure 1 show that referents of 3rd person pronouns display a higher discourse prominence than those of demonstratives. With personal pronouns, the number of cases in which the referent is not resumed in the following 10 sentences at all (0x) is much smaller (only 3 instances) than with demonstratives (11 instances). At the same time, referents of personal pronouns are resumed seven or more than seven times in the following context more often (14 instances) than referents of demonstratives (10 instances). Let us interpret the results for the middle columns. Here, we observe that personal pronouns less often resume a referent which will be mentioned only once in the following context (1x), but with two and more mentions (2x–6x), the picture changes, in that personal pronouns again correlate with more frequent mentions of the respective referent than demonstratives. This suggests that personal pronouns resume referents which are intended to be

present in the discourse in the long term, as major protagonists, while demonstratives are chosen to resume referents which are important only within a short discourse section, representing less relevant, minor characters.

Consider an illustrative minimal pair in (18) and (19):

(18) *uuas giuuortan gotes   uuort/ ubar* **iohannem[1] zachariases[2]**
     was done   God-GEN word about John-ACC Zacharias-GEN
     sun/ In thero   uuostinnu,/ Inti quam **her[1]** *In alle thie*
     son In the-DAT desert-DAT and came PERS.PR in all the-ACC.PL
     lantscaf/ Iordanis   predigonti toufi   riuua
     regions   Jordan-GEN preaching baptism repentance
     'The word of God came unto John the son of Zacharias in the wilderness and
     he came into all the country about Jordan preaching.'
     Lat. factum est verbum dei/super Iohannem zacharię filium/In deserto./Et
     uenit In omnem regionem/Iordanis praedicans baptismum pænitentię

     (T 43, 27–31)

(19) **Iusta[1]** *Iúuares   fater[2] uuoll& Ir   tuon. /* **ther** *uuas manslago*
     lusts   your-GEN father want   you do-Inf DEM was   murderer
     fon   anaginne.
     from beginning-DAT
     'The lusts of your father you will do. He was a murderer from the beginning.'
     Lat. & desiderata patris uestri uultis facere./ille homicida erat ab Initio.

     (T 218, 20–21)

In (18), the personal pronoun *her* takes up the more salient of two pre-mentioned referents, namely the one whose antecedent is realized within a directional PP selected by the main verb, in contrast to the less salient one realized as a genitive modifier. In (19), we observe the inverse relation, i.e. the demonstrative *ther* takes up the less salient of two referents, i.e. the one that is realized as the genitive modifier, in contrast to the more salient referent of the head of this modifier, the subject 'the lusts'. But the selection of the anaphora coincides with the role of the referents in the following discourse. While (18) introduces a lengthier passage in which John is a major protagonist and is perpetually resumed by anaphoric expressions, the referent of the demonstrative in (19) is not resumed in the following context because the discourse does not proceed on it.

It is revealing to compare at what stage in the development of the subsequent discourse the mentions of the referents of different anaphora occur. The relevant information is provided by Table 1. The leftmost column provides the number of sentence after resumption by the critical anaphor. The remaining columns present the numbers of mention of referents per number of clause and the percentage of these mentions for personal pronouns and demonstratives in the OHG corpus.

Table 1. Number of mentions of referents of personal pronouns and demonstrative per clause

Number of sentence	3rd person pronouns		demonstratives	
	total	%	total	%
clause 1	28	14.6%	24	16.7%
clause 2	25	13.0%	20	13.9%
clause 3	22	11.5%	22	15.3%
clause 4	21	10.9%	14	9.7%
clause 5	19	9.9%	17	11.8%
clause 6	17	8.9%	10	6.9%
clause 7	17	8.9%	11	7.6%
clause 8	15	7.8%	6	4.2%
clause 9	13	6.8%	10	6.9%
clause 10	15	7.8%	10	6.9%
Total	**192**	**100%**	**144**	**100%**

Looking at the total number of resumptions (bottom line), we recognize that referents taken up by demonstratives are generally less frequently mentioned in the following context than those taken up by personal pronouns. Additionally, we see that mentioning of a referent resumed by a demonstrative basically takes place within the first three clauses after first resumption by the critical anaphor. Referents of personal pronouns are also mentioned more frequently in the first few sentences, but their resumption is more balanced along the entire discourse span of 10 sentences. This confirms the suggestion that demonstratives resume discourse referents which remain activated only within a limited span of time in the discourse. This result is in line with the conclusion drawn from Figure 1 suggesting that these referents have a lower discourse persistence compared to personal pronouns.

3.3 Case study II: *her/ther/theser* vs. (*g*)*ener* in OHG

Forms of the forerunner of the modern German demonstrative *jener* are attested in historical records in contexts in which the differentiation between various referents or groups of referents is required. Behaghel (1923) provides valuable examples from different historical periods of German and some observations on the semantic properties of this demonstrative.

In general, Behaghel claims that the anaphoric function of *jener* is a recent development, which takes place as late as towards the Early New High German period. The earliest examples that Behaghel (1923) finds are from the MHG period, however, they are inconclusive, because *jener* fails to resume a referent of the

preceding discourse. Behaghel considers as an early instance of anaphoric *jener* the MHG example given in (20), but here, the context does not facilitate a clear identification of the antecedent:

(20) *wer sleht **den** lewen[1]? wer sleht **den** risen[2]? wer*
who kills the-ACC lion-ACC who kills the-ACC giant-ACC who
*überwindet **jenen**[1]/[2] unt **disen**[1]/[2]?*
defeat Dist.DEM-ACC and DEM-ACC
'Who will kill the lion? Who will kill the giant? Who will defeat the one and the other?' (Walth 81: 7; Behaghel 1923: I, 294–295)

Note, however, that Haferland (2013: 57) proposes a significantly different interpretation of this example. In his view, the demonstratives do not resume the lion and the giant in the previous context but refer to additional, not explicitly named enemies. Consequently, this example should be considered invalid as well.

A much more conclusive picture can be drawn from the examples that Behaghel (1923) cites from later periods of German, e.g. from Early New High German (21) and the transition towards Modern High German (22).[13] In both of the cases, *jener* takes up the referent that has been introduced first, i.e. *the good day* in (21) or *soil* in (22), while the demonstratives *diesen* and *die* resume the entity that has been introduced later, i.e. *the bad day* in (21) and *air* in (22):

(21) *am **guten** Tage[1] sey guter*
on = the-DAT.SG good-DAT.SG day-DAT.SG be-Imp good-GEN.PL
*dinge und **den** bösen tag[2] nim auch fur gut,*
things-GEN and the-ACC bad-ACC day-ACC take also as good
*denn **diesen**[2] schafft Gott neben **jenem**[1]*
because DEM-ACC creates God next DIST.DEM-DAT
'Be of good temper both on good days and on bad days because the latter ones God creates next to the former ones.'
(Luther Pred Sal. 7,15; Behaghel 1923, I, 295)

(22) *Es läßt sich übel paaren **die Erde**[1] zu **der** Luft[2],*
EXPL allows REFL badly match the soil to the-DAT air
*dann **die**[2] will oben fahren, und **jene**[1] sinkt in sich*
because DEM wants up go-INF and DIST.DEM falls in REFL
'It is impossible to match soil with air because the latter one wants to go up but the former one falls down.' (Logau I, 2,13; Behaghel 1923, I, 294)

---

13. Hafeland (2013) also argues that the anaphoric use of *jener* is a young development, which is not established before the Early New High German period.

From this, Behaghel concludes that the meaning of the anaphor *jener* is to resume a referent introduced prior to another one in the same context:

> "[…] in der Anaphora bezeichnet *der* die nähere, *jener* die fernere Größe" [in anaphoric function, *der* resumes the proximal, *jener* the remote entity]
>
> (Behaghel 1923: Vol. I, 294)

> "[…] *dieser* bezeichnet die nähere, *jener* die fernere Größe" [*dieser* resumes the proximal, *jener* the remote entity] (ibid.)

In other words, the principle that governs the distribution of *jener* in combination with other anaphora is that of the linear distance to the antecedent, with *jener* being the anaphor resuming the entity mentioned earlier in the discourse. This is exactly what according to reference grammars determines the anaphoric uses of *dieser* and *jener* in modern German:[14]

> "Das leicht archaisierende Demonstrativum *jener/jene/jenes* verweist im Gegensatz zu *dieser/diese/dieses* ursprünglich auf einen von der Sprecherin etwas weiter entfernteren Ort. Dieser Unterschied wird in Texten der Gegenwart noch genutzt, wenn auf zwei nacheinander genannte Gegenstände oder Personen Bezug genommen werden soll: *Old Shatterhand saß neben Winnetou, und während dieser schweigend zuhörte, berichtete ihm jener von den Ereignissen der letzten Tage.* Mit *dieser* wird dann der im Text näher liegende (also zuletzt genannte), mit *jener* der etwas weiter entfernte Bezugspunkt bezeichnet" [In contrast to *dieser/diese/dieses,* the slightly archaic demonstrative *jener/jene/jenes* originally refers to a place which is more remote from the speaker. This difference is still exploited in contemporary texts, whenever reference to two successively mentioned objects or persons needs to be established: *Old Shatterhand was seated next to Winnetou and while the latter one was listening quietly, the former one told him about the events of the last days.* Consequently, *dieser* resumes the closer (last mentioned) entity, *jener* the remote one] (Hentschel & Weydt 2003: 246)[15]

But against the picture drawn from the examples discussed above, Behaghel observes that the distribution of the proximal and distal demonstrative occasionally violates the principle of the linear distance to the antecedent. E.g., in (23), *jener* refers to the closer antecedent, namely to *Primrosens Sophie*, while *diese* takes up *Friederike*, the remote antecedent.

---

14. For this reason, I will use the terms 'proximal' and 'distal' demonstrative for *der/dieser* and *jener* respectively. The latter will be glossed DIST.DEM in the examples.

15. See also Helbig & Buscha (1987: 256).

(23) **Friederiken**[1] *an   die   Stelle von* **Primrosens   Sophie**[2] *zu*
      Friederike-ACC at   the-ACC place of   Primrose-GEN Sophie   to
      *setzen,   war nicht schwer: denn   von   **jener**[2]   ist wenig*
      put-INF   was not   difficult because about DIST.DEM-DAT is few
      *gesagt, man gibt   nur zu, daß **sie**[2]   liebenswürdig sey;*
      said   Indef admits only PRT that PERS.PR amiable   is-PERS.PR
      diese[1] *war es wirklich.*
      DEM   was it for sure

      'It was not difficult to replace Friederike by Sophie Primrose, because about
      the latter one, it is said that she was amiable, the former one was it for sure.'

                              (Goethe, DW, G XXVII, 354, 4; Behaghel 1923, I, 295;
                              Goethe, Werke, ed. Trunz and Blumenthal 2002: Vol. 9, 435)

Behaghel tries to explain this instance along the lines of the previous examples by
applying the notion of the "mental" distance to a referent, which overrides the linear
distance to the antecedent:

> "Die geistig entferntere Größe kann durch *jener* aufgenommen werden, auch wenn
> sie in der Wortstellung das Nähere ist" [The mentally remote entity may be resumed
> by *jener*, even if it is the closer one in terms of word order]
>
>   (Behaghel 1923: I, 295)

Behaghel does not provide a more explicit definition of his notion of *mental distance*. What is clear, however, is that he intends to associate remoteness with the
salience and importance of protagonists in the mental state of the interlocutors,
rather than with the linear order of the occurrence of their antecedents in the discourse. Applying this interpretation to the example in (23), we would hypothesize
that Sophie, the referent of the distal demonstrative *jener*, is less important in the
respective context than *Friederike*, the referent of the proximal demonstrative *diese*.

   I took Behaghel's notion of *mental distance* to be the correlate of discourse
prominence in terms of discourse potential, and checked the context of the respective example in Goethe's autobiography *From my Life: Poetry and Truth*. The
example is located in Book 10, referring to Goethe's journey to Sesenheim and
his encounter with Friederike Brion at her parents' home. Prior to this scene, the
Primrose family is described. But in the critical context, Sophie is subject to a singular, occasional mention and is completely dropped in the following discourse,
while *Friederike* is maintained as the major protagonist of the entire chapter.

   This raises the question if the distribution of the various types of demonstratives in historical German is governed by intended discourse prominence of the
referents rather than by the linear distance to the antecedent, with *jener* resuming
the less prominent protagonist in the global discourse. I searched the TITUS database to determine the earliest available examples and to analyze the discourse

properties of the respective antecedents. I found conclusive examples in texts from both OHG and MHG, see (24) – (25), suggesting that, contrary to the statements in the previous literature, *jener* appears in anaphoric function in combination with another anaphor very early in the attestation:

(24) **Ádaman　then　álton[1]** bisuéih　**er[3 = the Devil]** mit
　　　Adam-ACC the-ACC old-ACC appeased PERS.PR　　　with
　　　*then　　wórton, /　**ther júngo joh ther gúato[2]** giréh*
　　　the-DAT.PL words-DAT.PL the　young　and the　noble　avenged
　　　inan[1]　　gimúato. / *Spúan* **er[3]**　*io　zi nóti　***jénan[1]***
　　　PERS.PR-ACC properly　misled PERS.PR ever unfortunately DIST.DEM
　　　zi úbarmuati/[…] **Er[3]**　*wolta　in* **thémo[2]** *ana　wánk*
　　　to haughtiness-DAT PERS.PR wanted in DEM-DAT without doubt
　　　duan　so　samalichan skránk; **génan[1]**　*so*
　　　do-INF so same-ACC　baseness DIST.DEM-ACC that way
　　　bifált　**er[3]**
　　　destroyed he
　　　'He appeased the Old Adam with words. The young and noble one avenged him in a proper way. Prior to that, he [the devil] had unfortunately misled the former one [Adam] to haughtiness […], without a doubt, he [the devil] wanted to play the same base trick to the latter one [to Jesus], the former one [Adam], he had destroyed in that way.'　　　　　　　　　(O II, 5, 5–14)

(25) *dô　riet　er sâ /　daz　mir　bî den　zîten*
　　　then advised he therefore that me-DAT by the-DAT.PL times-DAT.PL
　　　dâ /　gezæme　wol **Dêmêtrîus** / *unde* **Pêkulâus[1]**/ *und* **dirre**
　　　then was proper well Demetrius　and　Pekulaus　and this
　　　Dimnus[2] *der　hie　lît, /　daz　sie　mîn　phlægen*
　　　Dimnus　who　here lies　that　they me-GEN take.care-INF.
　　　*zaller　zît. /　**disem[2]** einem　was ich gram, /　**jene***
　　　to = all-DAT time DEM-DAT one-DAT was I　aggrieved at DIST.DEM
　　　zwêne[1] *ich an mich　nam*
　　　two-ACC.PL I　to　me-ACC took
　　　'He advised then therefore that it was proper for me at that time that Demetrius and Pekulaus and this Dimnus who is laid there took care for me for all time ever. I was aggrieved at the latter, the former two I took to me.'
　　　　　　　　　(Rudolf von Ems, Alexander: Alex., 243, Rud., V Buch, 19362 ff.;
　　　　　　　　　　　　　　　　　　　　　　　　cit. in TITUS)

I will concentrate on the OHG data and investigate the referential properties of the different types of demonstratives in the examples that the corpus produced.

I found three examples of anaphoric pairs in Otfrid's *Gospel Book* and six[16] in Notker's works. In one example in Otfrid's work and in two examples from Notker, the distribution of *der/dieser* and *jener* confirms the principle of the linear distance to the antecedent, in that *jener* takes up the remote antecedent. Consider (26). The OHG equivalent of *jener* resumes the contents of the earlier utterance, namely that the convict confesses to having supported the infirmity of the Roman Senate. The demonstrative *tísses*, at the same time, resumes what the convict rejects, namely the contents of the closer utterance:[17]

(26)   *At*   *uolui*   *senatum*   *saluum*   *esse [...]*   *Dés*   *iího*   *ih.*   *Íh*
     and   wished   senat-ACC   sound-ACC   be-INF   this-GEN   confess   I   I
     neírta   *dóh*   *ten*   *méldare*   *nîeht. [...]*   ***Énes***
     NEG-impeded   though   the-ACC   messenger   NEG   DIST.DEM-ACC
     iího   *ih* .   ***tísses***   *neiího*   *íh*
     confess   I   DEM-ACC   NEG-confess   I
     'I wanted the Senate to be sound [...] This I confess [...]. Nevertheless, I did not
     impede the messenger. The former one I confess, the latter one I do not confess.'
                                                         (NBCons I, 24, 16–20)

But two examples from Otfrid and another four from Notker contradict this picture, in that *jener* resumes the closer antecedent. Consider the data from Otfrid's *Gospel Book* in (27) and (28) first:

(27)   *bi thiu*   *mág*   *er[1]*   *sin*   *in áhtu*   *théra*   ***Davídes[2]***
     therefore   can   PERS.PR   be-Inf   in   estimation-DAT   the-DAT   David-GEN
     slahtu. ...]   *Ríhta*   ***gener [2]***   *scóno*   *thie gótes*   *liuti*   *in fróno: so*
     lineage[...]   led   DIST.DEM   already   the   God-GEN   tribes   in   bliss   so
     duit ouh   ***thérer[1]***   *ubar jár*
     does   also   DEM   over   years
     'For this reason, he can be esteemed to be of the lineage of David. The latter
     already has led the people of God towards bliss, the former one is doing the
     same thing year over year.'   (O, Ad Lud., 56–60)

---

16. Actually, the corpus provided eight examples but I excluded two of them because the interpretation of the anaphora is ambiguous. These are examples (NPs 82, 306, 7) and (NPs 101, 374, 30). Resolving *ther/theser* and *ener* as taking up both antecedents reveals proper interpretations. In the first example, both those who return to Jesus and those who betrayed Him must confess that He is the superior God, and in the second one, both referents, the lower skies and the upper skies, will collapse and break into each other on doomsday.

17. The remaining examples are (O II, 5, 5–8) given in (25) and (NPs 1, 10, 15).

(28) **Pétrus[1]** ward   es ánawert joh bratt **er[1]**   slíumo thaz
      Peter   became it aware of and took PER.PR quickly the-ACC
      suert, /   **er[1]**   hérzen   sih gihárta   inti **éinan [2]** sar
      sword   PERS.PR courage-ACC REFL plucked up and one-ACC quickly
      irwárta. / Ih wéiz **er[1]**   thes   ouh fárta,   thes
      wounded I know PERS.PR Corr-GEN also intended the-GEN
      hóubites   rámta, / tház **er[1]**   thaz gisítoti,   then
      head-GEN hit-PRET for PERS.PR CORR achieved-PRET.SUBJ the
      méistar irrétiti; /   Gistuant **géner [2]** (wan ih) thénken,
      master rescued-PRET.SUBJ Started DIST.DEM guess I   think-INF.
      tház **er[2]**   wolti   wénken, / thoh   slúag **er[1]**
      that PERS.PR wanted-PRET.SUBJ escape   nevertheless cut PER.PR
      imo[2] in wára thana thaz   zésua   ora.
      PERS.PR indeed off   the-ACC right-ACC ear-ACC
      'Peter became aware of it and quickly fetched his sword. He plucked courage
      and injured one of them. I know, he intended to hit his head, in order to rescue
      his master. The other one thought, he would escape, but nevertheless he[Peter]
      cut him off the right ear.'   (O IV, 17, 1–6)

(27) is found at the very beginning of the *Gospel Book*, in the chapter dedicated
to King Louis the German, the ruler of Eastern Francia. The chapter glorifies the
braveness and piety of the Franconian ruler. In the critical context, he is compared
with King David, who according to the Old Testament shared and mastered similar
challenges during his reign. The distal demonstrative *gener* takes up King David,
although he is the closer antecedent, while the compound demonstrative *therer*
refers to King Louis, the remote antecedent. Similarly, in (28), in the context of
Jesus's arrestment in the Garden of Gethsemane, *gener* takes the referent introduced
later in the context, namely one of the soldiers. By contrast, the referent introduced
earlier in the context, Peter, is resumed by the personal pronoun *er*.

    Let us finally consider a representative example from Notker given in (29).
Here, *jener* takes up the closer referent, namely the Egyptians, the enemies of God's
people, while *dise* refers to the remote antecedent, the Israelites:[18]

---

18. The remaining examples are (NPs 45, 162, 4), (NPs 93, 349, 35) and (NBCons II, 105, 1).

(29) *Do   mêrota   Got **sînen**   **liût**[1] . uuanda er mánigfaltota*
then increased God his-ACC people because he multiplied

*in[1] .   unde starchta   in[1]   úber   **sîne**   **fienda**[2].*
PER.PR-ACC and   strenghted PER.PR-ACC against his-ACC enemies

*getéta in[1]   ôberoren **sînen**   **fienden**[2] . also in rubro*
made PERS.PR-ACC superior his-DAT.PL enemies-ACC as in red

*mari skeîn .   dô   **díse**[1] chamen ûz . unde **éne**[2]*
sea   happened when DEM-PL came   out and DIST.DEM

lágen ínne
laid within

'God increased his people as He multiplied it and made it strong against its
enemies and made it superior to his enemies, as became clear at the Red Sea,
when the former ones escaped and the latter ones remained inside.'

(NPs 104, 391, 24–392, 1)

In sum, the principle of the linear distance to the antecedent is much more often
violated than maintained in the OHG data. This leads us to the question if there is
an alternative mechanism that governs the distribution of the demonstratives in
the examples.

I hypothesized discourse prominence to be a factor influencing the selection
of anaphora and determined the frequency of explicit mention of the respective
referents in 10 subsequent clauses. Figure 2 provides the results.

Discourse prominence of referents of *her/ther/theser* and (*g*)*ener* in OHG
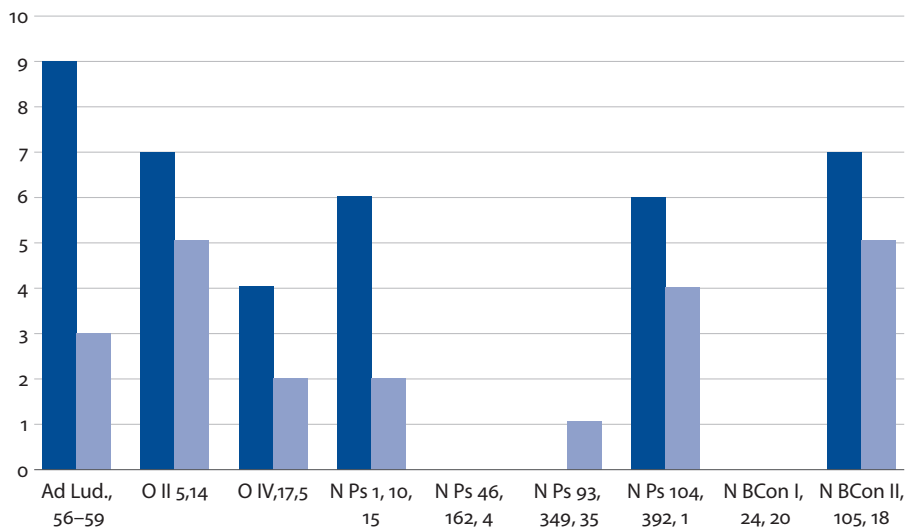
Figure 2. Number of resumption of referents of anaphoric pairs in 10 subsequent clauses
in OHG (total values)

For each example, the left-hand column represents the frequency of resumption of the referent picked up by *her/ther/theser*, while the right-hand one gives the respective numbers for the referent of *(g)ener*. The results show that with the exception of one example (NPs 93, 349, 35), the referent picked up by *jener* is resumed less often in the following context. This means that compared to the referent of the alternative anaphoric expression that occurs in the same context, the referent resumed by the distal demonstrative *jener* in OHG displays a lower degree of discourse prominence, irrespective of the activation state of this referent in terms of order of mention.

Another important result can be observed from the data. It pertains to discourse prominence in terms of protagonistship, and is especially well represented in Otfrid's work which is prototypically narrative, in contrast to the theological and exegetic style of Notker's work. In none of Otfrid's examples does *jener* refer to the major protagonist of the respective discourse section. It is rather its alternative that takes up the more important character, e.g. Jesus, who is the main character in the scene about the temptations in the desert (24) or King Louis, who is the subject of glorification in (27). This means that the distal pronoun *jener* is systematically used to signal that the referent that is resumed by it plays a minor, less important role in the discourse.

4. Conclusion

Starting from the observation that anaphoric reference is a major device of creating coherence in narrative discourse, the present paper investigates the referent's discourse prominence as a factor in anaphora selection. On the basis of data from historical German, it determines the discourse persistence of referents resumed by various types of anaphora, measured in terms of frequency of explicit mention in a series of subsequent clauses. Two case studies are presented. The first one examines the seemingly free variation in the anaphoric use of 3rd person pronouns vs. simple demonstratives in OHG, the second one deals with the principles of the distribution of different types of demonstratives resuming distinct antecedents in the same context. Both case studies confirmed that the referent's intended discourse prominence has an impact on the choice of the anaphor with which it is resumed for the first time after introduction by a lexical DP. More precisely, it was shown that anaphora with a lower degree of lexical explicitness are used to resume referents which are taken up more frequently and within a longer span in the following context, while anaphora with a higher degree of explicitness correlate with a lower degree of persistence of the referent in the following context.

This outcome does not only contribute to a better understanding of the distribution and the semantic properties of different lexical types of pronouns in historical

German but also provides empirical evidence in favor of the view that the form of a nominal expression depends on the properties of the referent in the subsequent discourse. At present, discourse persistence is a well-established factor in research on variation between different semantic types of nominal expressions, including their syntactic realization, but its role in anaphora choice is still unclear. An insightful suggestion is put forward by Garnham and Cowle (2008), who account for a remarkable Janus-like behavior of pronouns in interpretation. The present paper confirms this view by showing that the choice of anaphoric expressions is not only determined by the properties of the antecedent in the previous discourse but also by those in the subsequent context.

References

Text editions

MF *The Monsee Fragments. Newly collated text with introduction, grammatical treatise and exhaustive glossary, and a photo-litographic fac-simile*, ed George Allison Hench, Strassburg: Trübner. 1890.

NBCons *Die Werke Notkers des Deutschen*. Vol. 1–2. *Boethius, De consolatione philosophiae* Books I/II and III/IV, ed Petrus W. Tax. Tübingen: Niemeyer. 1986–1988.

NPs *Die Werke Notkers des Deutschen*. Vol. 8–10. *Der Psalter*, ed Petrus W. Tax. Tübingen: Niemeyer. 1979–1983.

O *Otfrids Evangelienbuch*, ed Oskar Erdmann, Tübingen: Niemeyer. 1973.

T *Die lateinisch-althochdeutsche Tatianbilingue Stiftbibliothek St. Gallen Cod. 56*, ed Achim Masser, Göttingen: Vandenhœk & Ruprecht. 1994.

Johann Wolfgang von Goethe, Werke, Kommentare und Register. Hamburger Ausgabe in 14 Bänden, ed Erich Trunz and Lieselotte Blumenthal. Band 9. *Autobiographische Schriften I*. 14th ed. München: Beck. 2002.

Secondary literature

Abraham, W. 2002. Pronomina im Diskurs: Deutsche Personal- und Demonstrativpronomina unter 'Zentrierungsperspektive'. Grammatische Überlegungen zu einer Teiltheorie der Textkohärenz. *Sprachwissenschaft* 27: 447–491.

Almor, A. & Nair, V. A. 2007. The form of referential expressions in discourse. *Language and Linguistics Compass* 1(1–2): 84–99.

Ariel, M. 1988. Referring and accessibility. *Journal of Linguistics* 24: 65–87.

Ariel, M. 2001. Accessibility theory: An overview. In *Text Representation: Linguistic and Psycholinguistic Aspects* [Human Cognitive Processing 8], T. Sanders, J. Schilperoord, & W. Spooren (eds), 29–87. Amsterdam: John Benjamins.

Arnold, J. E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4: 187–203.

Asher, N. 1993. *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer.

Axel, K. & Weiß, H. 2011. Pro-drop in the history of German. From Old High German to the modern dialects. In *Empty Pronouns*, P. Gallmann, & M. Wratil (eds), 21–51. Berlin: De Gruyter.

Axel, K. 2007. *Studies in Old High German Syntax: Left Sentence Periphery, Verb Placement and Verb-second* [Linguistik Aktuell/Linguistics Today 112]. Amsterdam: John Benjamins.

Behaghel, O. 1923. *Deutsche Syntax. Eine geschichtliche Darstellung*, Vol. I: *Die Wortklassen und Wortformen. A. Nomen. Pronomen*. Heidelberg: Winter.

Bildhauer, F. & Cook, P. 2010. German multiple fronting and expected topic hood. In *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar*, 68–79. Stanford CA: CSLI.

Bosch, P. & Umbach, C. 2007. Reference determination for demonstrative pronouns. *ZAS Papers in Linguistics* 48: 39–51.

Bosch, P., Katz, G. & Umbach, C. 2007. The non-subject bias of German demonstrative pronouns. In *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference* [Studies in Language Companion Series 86], M. Schwarz-Friesel, M. Consten & M. Knees (eds), 145–164. Amsterdam: John Benjamins.

Bosch, P., Rozario, T. & Zhao, Y. 2003. Demonstrative pronouns and personal pronouns. German *der* vs. *er*. In *Proceedings of the EACL Workshop on the Computational Treatment of Anaphora in Budapest 2003*. Stroudsberg PA: ACL. http://cogsci.uni-osnabrueck.de/~pbosch/download/bosch2003_tr&yz_demonstratives.pdf

Braune, W. 2004. *Althochdeutsche Grammatik,* Vol. I: *Laut- und Formenlehre*, 15th edn. Tübingen: Niemeyer.

Brennan, S. E., Friedman, M. W. & Pollard, C. J. 1987. A centering approach to pronouns. In *Proceedings of the 25th annual meeting of the Association for Computational Linguistics, July 06–09, Stanford, California*, 155–162. Stroudsberg PA: ACL.

Cozijn, R., Commandeur, E., Vonk, W. & Noordman, L. G. M. 2011. The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language* 64(4): 381–403.

Deichsel, A. 2011. The discourse effects of the indefinite demonstrative *dieser* in German. In *Proceedings of the 23rd European Summer School in Logic, Language, and Information (ESSLLI), 1st–12th August 2011, Ljubljana, Slovenia*. 70–77. <http://www.stanford.edu/~danlass/esslli2011stus/proceedings.html>

Deichsel, A. & von Heusinger, K. 2011. The cataphoric potential of indefinites in German. In *Anaphora and Reference Resolution. 8th Discourse Anaphora and Anaphor Resolution Colloquium*, I. Henrickx, S. L. Devi, A. Branco & R. Mitkov (eds), 144–156. Heidelberg: Springer.

Demske, U. 2005. Weshalb Demonstrativpronomina nicht immer Determinantien sind. In *Deutsche Syntax: Empirie und Theorie*, F.-J. d'Avis (ed.), 53–80. Göteborg: Göteborger Germanistische Forschungen.

Duden. 1995. *Die Grammatik*, 5th edn. Mannheim: Dudenverlag.

Garnham, A. & Cowles, H. W. 2008. Looking both ways. The JANUS model of noun phrase anaphor processing. In *Reference. Interdisciplinary Perspectives*, J. Gundel & N. Hedberg (eds), 246–272. Oxford: OUP.

Garnham, A., Traxler, M., Oakhill, J. & Gernsbacher, M. A. 1996. The locus of implicit causality effects in comprehension. *Journal of Memory and Language* 35(4): 517–543.

Givón, T. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 2(2): 35–53.

Givón, T. 1983. Topic continuity in discourse. An introduction. In *Topic Continuity in Discourse. A Quantitative Cross-language Study* [Typological Studies in Language 3], T. Givón (ed.), 1–41. Amsterdam: John Benjamins.

Grüning, A. & Kibrik, A. A. 2005. Modelling referential choice in discourse: A cognitive calculative approach and a neutral network approach, In *Anaphora Processing. Linguistic, Cognitive and Computations Modelling* [Current Issues in Linguistic Theory 263], A. Branco, T. McEnery & R. Mitkov (eds), 163–197. Amsterdam: John Benjamins.

Gundel, J. K., Hedberg, N. & Zacharsky, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2): 274–307.

Haferland, H. 2013. Vokale Kultur, Hörgedächtnis und Textgrammatik. Zur Pronominalisierung in mittelhochdeutschen Texten. In *'der âventiuren dôn': Klang, Hören und Hörgemeinschaften in der deutschen Literatur des Mittelalters*, I. Bennewitz & W. Layher (eds), 45–62. Wiesbaden: Reichert.

Helbig, G. & Buscha, J. 1987. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*, 10th edn. Leipzig: Enzyklopädie.

Hentschel, E. & Weydt, H. 2003. *Handbuch der deutschen Grammatik*, 3rd edn. Berlin: De Gruyter.

Ionin, T. 2006. This is definitely specific: Specificity and definiteness in article systems. *Natural Language Semantics* 14: 175–234.

Ionin, T. 2013. Pragmatic variation among specificity markers. In *Different Kinds of Specificity across Languages*, C. Ebert & S. Hinterwimmer (eds), 75–104. Heidelberg: Springer.

Petrova, S. & Solf, M. 2010. Pronominale Wiederaufnahme im älteren Deutsch. Personal- vs. Demonstrativpronomen im Althochdeutschen. In *Historische Textgrammatik und historische Syntax des Deutschen. Traditionen, Innovationen, Perspektiven*, A. Ziegler & C. Braun (eds), 339–364. Berlin: De Gruyter.

Prince, E. F. 1981. Toward a taxonomy of given-new information. In *Radical Pragmatics*, P. Cole (ed.), 223–255. New York NY: Academic Press.

Solf, M. 2008. Status und Zugänglichkeit von Diskursreferenten im Althochdeutschen am Beispiel der Tatianbilinue Cod. Sang. 56. In *Die Formen der Wiederaufnahme im älteren Deutschen*, Y. Desportes, F. Simmler & C. Wich-Reif (eds), 177–197. Berlin: Weidler.

Strube, M. & Hahn, U. 1999. Functional centering – Grounding referential coherence in information structure. *Computational Linguistics* 25(3): 309–344.

Vinckel, H. 2006. *Die diskursstrategische Bedeutung des Nachfelds im Deutschen. Eine Untersuchung anhand politischer Reden der Gegenwartssprache*. Wiesbaden: Deutscher Universitätsverlag.

Vinckel-Roisin, H. 2011. Wortstellungsvariation und Salienz von Diskursreferenten. Die Besetzung des Nachfeldes in deutschen Pressetexten als kohärenzstiftendes Mittel. *Zeitschrift für Germanistische Linguistik* 39: 377–404.

Wright, S. & Givón, T. 1987. The pragmatics of indefinite reference: Quantified text-based studies. *Studies in Language* 11(1): 1–33.

A centering theoretic account
for the changing usage of anaphoric
expressions in the history of German

Augustin Speyer

Saarland University

The Centering status of a referent and the coherence relation of the discourse
unit in which the referent is used are potential factors for anaphor resolution.
In Old High German, as exemplified with Otfrid's Evangelienbuch, there is a
strong tendency for this generalisation – Backward looking centers in Continue
relations tend to be realised by personal pronouns, in Retain and Shift relations,
they tend to be realised by demonstrative pronouns – but it is far from categor-
ical. There are even blatant violations of generalizations that hold for Modern
German, so the importance of Centering as a factor is higher today than in Old
High German. In Early New High German of the 14th century, anaphora res-
olution was sensitive to Centering. The hierarchical structure of a text played a
bigger role, as adjacency was defined not strictly linearly, but with respect to dis-
course units on the same hierarchical level.

1. Introduction

1.1   Aims of this paper

An important cohesive devise in narrative texts is the usage of anaphoric expres-
sions. The thematic structure of narrative texts tends to be structured in such a way
that a limited number of referents is used as topics (with usually one main actor),
thus constructing the skeleton of the text from a thematic point of view, to which
additional, local topics may be added. By using anaphoric expressions the author
(or speaker) signals that the referent s/he is referring to at the moment is part of this
thematic structure; by selecting the type of anaphoric expression the author can give
the recipient a hint how prominent in the thematic structure the referent is. This is a
constant of telling stories and probably has been so forever. What changes, however,
is the set of anaphoric expressions an author can make use of, so that the association

of certain types of anaphoric expressions to certain degrees of prominence has constantly to be newly defined. This is what this paper is about, more concretely, the change in usage of anaphoric expressions in Old High German (OHG) and Early New High German (ENHG), viewed from the perspective of Centering Theory.

The study of anaphoric expressions, or to be more precise, of the question under what circumstances what expression may be used, was originally induced by seminal work such as Ariel (1990) or Gundel, Hedberg & Zacharski (1993), but gained enormous momentum in the past decade (cf. e.g. the papers in Stark, Leiss & Abraham 2007). A subfield of particular interest is the question of the historic development. This question has also been investigated thoroughly since the last couple years, a fact of which the papers collected in this volume give impressive testimony.

The history of German is of particular interest for that matter. We know that German was originally a pro-drop language, that OHG from its earliest attestations on is only a partial pro-drop language (e.g. Axel 2007; Axel & Weiß 2011), and that in the course of OHG (roughly from 750 to 1050) the modern system evolves, in which zero subject pronouns are no longer acceptable and in which two expressions are available for anaphoric, non-deictic reference, namely the personal pronoun (PerPr: *er, sie, es* etc.) and the simple demonstrative pronoun (DemPr: *der, die, das* etc.). Important questions that permeate the research on that subject are, what are the conditions under which any of the three available forms (zero, PerPr and DemPr) was used in OHG, how did these conditions change in the aftermath of the loss of the zero pronoun option, and are there any changes in usage independent of that matter. This opens up an extremely large field; there are excellent studies attempting to answer at least some of these questions (e.g. Schlachter 2012 and references there). In this study I am concentrating on a tiny lot on that field, mainly ignoring the questions of zero pronouns, phrasal ellipsis, full name repetition, instead trying to trace back the usage of PerPr and DemPr throughout the history of German – that is, confining myself to two time spans, OHG and ENHG (roughly 1350 to 1650) – in their anaphoric, non-deictic usage. I am doing this by using Centering Theory (CT; Grosz, Joshi & Weinstein 1995; Walker, Joshi & Prince 1998), as the aim of this theory is to measure textual coherence by tracing the salience of discourse referents. CT makes certain predictions about the relative salience of those referents, and combined with theories on the realization of different degrees of salience in terms of linguistic expressions, such as Ariel (1990), more worked out in (Abraham 2007a: 33f.), it also makes predictions when what forms are used. These predictions can be tested at the historic evidence.

In order to do that it is best first to give a brief overview of the main tenets of CT, as far as they are relevant here (Section 1.2). In Section 2, the usage of anaphors in Modern German is viewed under a CT perspective. This is applied first to OHG (Section 3) and then to ENHG (Section 4). It will become clear that CT makes best

predictions about ENHG, and that in fact the ENHG evidence may help to solve some problems that CT in general, but especially in its application to German (on that see Speyer 2007) is facing.

1.2   Centering theory

The main tenets of Centering Theory can be summarised as follows (see Grosz, Joshi & Weinstein 1995; Walker, Joshi & Prince 1998):

- All members of the set of referring expressions R of an utterance U_{n-1} have the potential to be resumed in the following utterance U_n. In CT terminology these referring expressions have by that the rank of **forward-looking centers (Cf)**.
- One element of $R(U_{n-1})$ has a higher likelihood to be resumed in U_n. In CT terminology, this referring expression is called **preferred center (Cp)**.
- At least one element of $R(U_n)$ is coreferential with one of the Cfs (most likely the Cp). This element that links U_n back to U_{n-1} and by that to the previous discourse is called **backward-looking center (Cb)** in CT terminology. The Cb is the element that constitutes textual coherence on a semantic level.

What determines the higher likelihood of a Cp to be resumed is a matter of debate and probably language specific. There have been a number of suggestions for different languages, ranging from the grammatical function in English (assuming a hierarchy *Subject > Object(s) > other referent*, the Cp is the expression with the highest rank in that hierarchy; see Grosz, Joshi & Weinstein 1995) to a hierarchy of thematic roles in Turkish (*Origo > Agens > Experiencer > Thema*; see Turan 1998).[1] For German, the ranking is unclear. A ranking similar to English is at least feasible (see e.g. Abraham 2007a: 35). A proposal was made by Strube & Hahn (1996) that the ranking depends on the topichood; in a similar vein go Bethke (1990) and Abraham (2002). In Speyer (2007), this view was criticised as being circular; if the goal is to model topichood as directly correlated to centerhood, any notions of topic (e.g. aboutness, givenness) should not be included in the definition of center, of which the Cf-ranking is an important part. One of the goals of this paper is to find evidence for the Cf-ranking that fits best the German data. Obviously, this is a question that cannot be solved here – in fact, it would make a full-fledged research project on its own – but the ENHG data is highly suggestive in that respect and might point to the right direction.

One of the main goals of CT is to model local discourse coherence. This is done by cross-classifying two parameters related to the backward-looking center of a

---

1.   These two options might be related, as grammatical functions are dependent on thematic roles (see e.g. Primus 2011).

given utterance, $Cb(U_n)$. The first parameter is, whether $Cb(U_n)$ is identical to the preferred center of the same utterance, $Cp(U_n)$, or not. This parameter has to do with the formal realization of the backward-looking center in the utterance itself, whether it is realised prominently or not. The second parameter is, whether $Cb(U_n)$ is identical to the backward-looking center of the preceding utterance, $Cb(U_{n-1})$. This parameter focuses more on the thematic progression, that is, whether a text passage sticks to a topic or changes it. There are four possible combinations of the values of these parameters, and these combinations define four coherence relations in CT, as illustrated in Table 1 (see Walker, Joshi & Prince 1998).

Table 1. Definition of coherence relations in centering theory

	$Cb(U_n) = Cb(U_{n-1})$	$Cb(U_n) \neq Cb(U_{n-1})$
$Cb(U_n) = Cp(U_n)$	Continue	Smooth shift
$Cb(U_n) \neq Cp(U_n)$	Retain	Rough shift

What is the relationship between Centering Theory and the study of anaphoric expressions? Centering Theory is, among other things, a theory of the accessibility of referents.[2] The accessibility is directly at issue when it comes to the ranking of forward-looking centers: The Cp can be understood as the most accessible referent, and it is this property that determines its rank. A text is maximally coherent if the most accessible referent is used over and over again – this is the Continue relation. A text looses coherence if a referent is chosen as center that is not the most accessible one (Retain) or if another referent from the preceding context is presented such that it is the most accessible referent (Smooth Shift) or if another referent is chosen as center that is not the most accessible one (Rough Shift). But accessibility is also an issue when it comes to the selection of anaphoric expressions, in that highly accessible referents are expressed by a given type of expression A_1, while less accessible referents are expressed by a given type A_2. So there is a direct link between Centering Theory and the selection of anaphoric pronouns; ideally it would be like in (1):

(1)   If a language has several types of anaphoric expressions $A_1, ..., A_n$, and if they are aligned with accessibility of the referent such that A_1 is the most accessible and A_n the least accessible, A_1 should be used preferably in utterances being in a highly coherent relationship to the preceding context (e.g. Continue), and the other expressions in descending order for less coherent relationship, also in descending order.

---

2.   Centering Theory is, of course, somewhat simplifying, in that e.g. memory decay or salience gradation is not modelled there. See on memory decay e.g. van Dyke 2012.

As there are other factors that influence accessibility than centerhood, the link in (1) will probably never show up in its purest form, but it should be noticeable as a tendency in the use of different types of anaphoric expressions.

2.   Anaphoric expressions in modern German from a centering perspective

In Modern German, there are basically two types of anaphoric expressions that might be subject to the link in (1), namely the personal pronoun (PerPr) *er, sie, es*, and the simple demonstrative pronouns (DemPr) *der, die, das*.[3] The latter developed into the definite article, the process of which is discussed thoroughly in Abraham (2007b).[4] Apart from that, there are other types of anaphoric expressions (such as the demonstrative pronouns *dieser, diese, dies(es)* or *jener, jene, jenes*), but they serve preferably deictic functions. I leave it open here whether they are subject to the accessibility hierarchy or the link given in (1); the research has been concentrating on the first two types of anaphoric expressions anyway.

There are a number of generalizations that have been made on the subject of anaphoric expressions in Modern German.[5] Some are related to the syntactic function, some to topichood. A generalization about the relation between the choice of anaphoric expressions and syntactic functions is, for instance, that PerPr preferably refer to subject antecedents, while DemPr preferably refer to non-subject antecedents (Bosch et al. 2003; Bosch et al. 2007). I will refer to this as *Subject Generalization*. A generalization about topichood and the choice of anaphoric expressions is such that PerPr preferably refer to discourse topics, while DemPr preferably refer to referents that are not discourse topics (Bosch et al. 2003; Bosch & Umbach 2007; Bosch 2012). I will refer to this as *Topic Generalization*. The latter has been discussed at length in Abraham (2007a). Both generalizations are presented in schematic form in Table 2.

---

3.   The story is actually more complicated than that; within PerPr there can be distinguished three degrees of 'strength' (clitic, weak, strong). The disstribution of DemPr shows certain additional differences to the one of PerPr (cf. e.g. Cardinaletti & Starke 1996: 55f.).

4.   As we are only dealing here with the 'absolute' uses of the pronoun – that is: uses in which it does not co-occur with an NP, but forms a DP by itself – I do not pursue further the interesting discussion about the grammaticalisation of the article out of the simple demonstrative pronoun laid out in Abraham (2007b).

5.   I am talking only about usages in which there is a clear antecedent. I am aware that this is only a part of the problem of pronominal usage (see Bosch & Umbach 2007 for discussion).

Table 2. Generalizations about the choice of anaphoric expressions in modern German

	grammatical function of antecedent	topichood of referent
PerPr	subject	discourse topic
DemPr	non-subject	not discourse topic

Note that both generalizations are compatible with generalizations that have been made in the framework of Centering Theory (s. Abraham 2002: 458f.). The Subject Generalization reminds of the Cf-ranking according to a scale of grammatical functions. If the Cfs are ranked according to grammatical functions, the Cp would be the subject under normal circumstances, and thus the Subject Generalization could be re-read as: PerPr refer to Cps, DemPr to referents that are not the Cp. The Topic Generalization underlies the generalization that PerPr correspond to the resumption of thematic or given information of the preceding sentence, while DemPr correspond to the resumption of rhematic or new information of the preceding sentence. In Abraham's (2007a) terminology, PerPr are assigned the function of *thema continuant*, while DemPr (which in Abraham are subdivided in 'real' demonstrative pronouns such as *dieser,-e,-es* and the simple demonstratives *der, die, das*) are termed *thema switchers*. In other words, PerPr are associated with the coherence relations Continue and Retain, whereas DemPr are associated with the relations Smooth and Rough Shift. The Topic Generalization makes an even stronger statement: DemPr could never refer to the $Cb(U_n)$ in a Continue or Retain relation, when they refer to a piece of new information in U_{n-1}, as it is impossible for new information to be a Cb. Note that the property of being a backward-looking center does not automatically involve that the referent is also the aboutness topic. A backward-looking center however fulfils *per definitionem* properties such as givenness and high salience ('thematicity') that are typical secondary properties of aboutness-topics. Aboutness can be seen as an interpretatory effect: The Cb, being the most accessible element, is preferably interpreted as the entity the sentence is about.

The Topic and Subject Generalizations in their Centering forms are illustrated in the sentence pairs (2).[6] All sentences in (2) are to be read under the assumption that Karl was mentioned in the preceding discourse, while the burglar was not.

---

6.   What follows is a somewhat simplified account of the relationship between centering status and the form of the anaphor. For a more detailed discussion, the reader is referred to Abraham (2007a). Note that Abraham, while using Centering concepts, focuses more on the thematic-rhematic properties that are hinted at by the centering relations.

(2) A: *[Karl]$_i$ sieht [einen Einbrecher]$_j$.*
     Karl sees a   burglar

   B: *[Er]$_{i/*j}$ gibt ein paar Warnschüsse ab.*
     PerPr gives a few warning shoots ptc.

   A: *[Karl]$_i$ sieht [einen Einbrecher]$_j$.*
     Karl sees a   burglar

   B: *[Der]$_{*i/j}$ gibt ein paar Warnschüsse ab.*
     DemPr gives a few warning shoots ptc.

   'Karl sees a burglar. He / This guy fires a few times for warning.'

   A: *[Karl]$_i$ sieht [einen Einbrecher]$_j$.*
     Karl sees a   burglar

   B: *Ein Krampf hindert [ihn]$_{i/?j}$ am Rennen.*
     A cramp hinders PerPr at running.

   A: *[Karl]$_i$ sieht [einen Einbrecher]$_j$.*
     Karl sees a   burglar

   B: *Ein Krampf hindert [den]$_{*i/j}$ am Rennen.*
     A cramp hinders DemPr at running.

   'Karl sees a burglar. A cramp handicaps him/this guy in his running.'

In (2a), the backward-looking center of sentence B (Cb$_B$) is Karl, the preferred center of sentence B (Cp$_B$) likewise is Karl. Karl is also the backward-looking center of the preceding sentence A (Cb$_A$), under the assumption that Karl is given in the preceding discourse. So this is an example of a Continue relation. If Cb$_B$ is Karl, only the PerPr can be chosen felicitously for reference. The PerPr cannot be interpreted as referring to a referent (such as 'einen Einbrecher') that is not the Cb$_B$ and not the Cp$_B$.

In (2b), the situation is different. The sentence is identical to (2a), save for the detail that the reference in B is made by a DemPr and not a PerPr. The Cb$_B$ is the burglar, as it is the only given referent in B. The Cp$_B$ is likewise the burglar, as it is in subject function. The Cb$_A$ however is still Karl, under the assumption that Karl is given in the preceding discourse. So this is an example of a Smooth Shift relation. In this case, reference is only possible by a demonstrative pronoun. From another perspective one could say: If in a situation as illustrated in (2a), (b) the subject of B is pronominalised and it is in form of a PerPr, it is automatically understood as a Continue relation. If it is in the form of a DemPr, it is spontaneously understood as a Smooth shift relation.

Example (2c) illustrates a Retain relation. The Cb$_B$ is Karl, but it is not the Cp$_B$, as it is not the subject in B. The subject, a cramp, is a new entity. The Cb$_A$ is again Karl. In this situation, reference by means of a PerPr can be interpreted as referring to Cb$_A$, but the PerPr could in theory also refer to the burglar. Note that the accessibility is not as high as in a Continue case.

If (2c) is changed such that reference is made by a DemPr, we get (2d). This sentence is understood automatically such that the DemPr refers to the Cb_B, which is the burglar. It must be the burglar, as this is the only given referent in (2dB). The Cb_B is however not the Cp_B which is the new entity 'a cramp', and it is not the Cb_A which is Karl, under the assumption that Karl is given in the preceding discourse and the burglar is not. So this illustrates a Rough Shift relation.

The examples in (2) illustrate that the form of the anaphoric expression is directly correlated to the coherence relation. The realization of the Cb by a PerPr is only compatible with Continue and Retain (with the proviso that in a Retain relation the PerPr need not refer to the Cb_A), the realization of the Cb by a DemPr is only compatible with Shifts.

From this property of DemPr to be associated with Shifts follows that theme continuation cannot be done by DemPr, especially if theme has already been resumed by PerPr (3; s. Abraham 2002: 461, Abraham 2007a: 32f., 38ff.).

(3) A: *[Karl]$_i$*   *sieht*   *[einen*   *Einbrecher]$_j$.*
        Karl   sees   a   burglar
     B: *[Er]$_{i/*j}$*   *gibt*   *ein*   *paar*   *Warnschüsse*   *ab.*
        PerPr   gives   a   few   warning shoots   ptc.
     C: *Er$_i$ / *Der$_i$*   *rennt*   *daraufhin*   *los.*
        he / this guy   runs   then   ptc.
     'Karl sees a burglar. He fires a few times for warning. Then he starts to run.'

Note that *der* in C of (3) could also refer to the burglar. This shows that a DemPr needs not refer to a center at all.

So we can say that the choice of an anaphoric expression is sensitive to the Centering status of the referent in Modern German. If it is a backward-looking center, the choice governs the interpretation such that PerPr are interpreted as referring to the entity that was the backward-looking center in the preceding sentence, indicating a Continue or Retain relation, whereas PerPr are interpreted as referring to an entity that was not the backward-looking center of the preceding sentence, indicating a Shift relation. So we see the linking given in (1) in operation: PerPr, which are higher on the accessibility hierarchy, are associated with more coherent relations (Continue and Retain), DemPr, which are lower, are associated with the less coherent Shift relations. Winter (2003) goes one step further in her hypothesis that DemPr select referents that lead to a less coherent coherence relation, whereas PerPr select referents that lead to a more coherent coherence relation. In Example (2a),(b), for instance, the subject of the B clause could either be Karl – leading to a Continue relation – or the burglar – leading to a Smooth Shift relation. As the coherence relations are ordered with respect to their coherence as follows (cf. Winter 2003: 13):

Continue > Retain > Smooth Shift > Rough Shift

Smooth Shift is lower on the hierarchy and therefore associated with the DemPr, whereas Continue is higher and thus associated with the PerPr.

3.   Old High German – does centering play a role for the choice of anaphoric expressions?

For OHG, generalizations have been made about the choice of anaphoric expressions. Petrova and Solf (2010) link the two variants PerPr and DemPr to salience: PerPr are used for more salient referents, such as the aboutness topic of the preceding discourse, whereas DemPr are used for less salient referents.[7] Salience can be defined as the degree of relative prominence of a unit of information (cf. Chiarcos, Claus & Grabski 2011). Factors influencing the salience of a referent are manifold, but among them are the grammatical function and topichood (Bosch and Umbach 2007), the same factors that we encountered already as determinative factors for the choice of anaphoric expressions in Modern German. So there is the possibility that modeling in CT could work for OHG as well.

In order to test this, the following two methods were chosen. The first method was a search for selected forms of DemPr, using the TITUS database. Starting from this data set the CT properties of the DemPr that were found was classified. The second method, targeting the PerPr (to have some point of comparison), was to make CT analyses of selected passages from the *Evangelienbuch* by Otfrid von Weißenburg (Alsace, around 870 AD).[8] Here the passages were selected randomly, but passages in which direct speech was involved were not taken into consideration, simply because there we would have to deal with origo shift in addition to anything else.

A question one might ask is: Why are we using a poetry text like Otfrid for a study like this? Normally, when doing syntactic research, Otfrid is treated with a good deal of caution, as the organization of the text in verses might be responsible

---

7.   Zero pronouns are not included here, although they would signal an even higher salience, following Ariel (1990). It is a matter of debate how much of a pro-drop language OHG still was (see for contrasting views e.g. Axel 2007; Schlachter 2012). Note that in the Tatian (which is a direct translation from a text written in the pro-drop language Latin) zero pronouns are much more frequent than in Otfrid (which is no translation). In fact, the usage of zero pronoun in Otfrid is almost as restricted as in Modern German (although there are some differences in usage, see Speyer 2016).

8.   The passages were Chapter 2,8; 2,9; 3,8;3,9; 4,2; 4,3; 5,5.

for a lot of weird syntactic phenomena that are not representative of ordinary 9th century OHG.

The answer is simple: The pros outweigh the contras. On the pro side, we have the fact that Otfrid is the only large original text we have from 9th century OHG. Translation texts are to be treated with caution, especially when dealing with anaphoric expressions. Latin, as is well known, was a pro-drop language; the set of anaphoric expressions was basically zero versus a bunch of demonstrative pronouns (*is, ea, id; ille, illa, illud; hic, haec, hoc* etc.) none of which is compatible with the PerPr of German. What the translators could have done would be to equate a Latin zero pronoun to a German PerPr, and a Latin DemPr to a German DemPr. This would be not very helpful for us, as this would simply replicate the Latin system. But they did not: They equated Latin zero pronouns sometimes with German zero pronouns and sometimes with German PerPr. It is hard to find a system behind the variation; opinions range from seeing zero pronouns in these translation texts as simply copied from Latin, where sometimes a PerPr slipped in at random (Eggenberger 1961), to a more systematic view: Axel (2007) observes that zero pronouns tend to occur postverbally in matrix clauses, Schlachter (2012) in her study on the Isidor group correlates zero pronouns to coordinative discourse relations. At any rate, using a translation text for a study on anaphoric pronouns encounters many problems.

On the contra side, we have the fact that Otfrid is no prose. But a closer look tells us that the verse organization of Otfrid should have no effect on the choice of anaphoric expressions. Why is that? First, PerPr and DemPr behave similarly from a metrical point of view: Virtually all forms of either pronoun are monosyllabic and unstressed (at least in anaphoric usage). Second, anaphoric expressions are usually not clause-final which makes them uninteresting as rhyme words (given that a verse equals a syntactic unit such as a clause, which in Otfrid is most often the case). Third, even zero pronoun as an alternative would make no difference in the verse of Otfrid, as the verse is composed with 'freie Senkung' (which means that a variable number of unstressed syllables can intervene between the stressed syllables; only the stressed syllables count), so the choice between zero pronoun and overt pronoun, be it a PerPr or a DemPr, is independent of metric considerations. This makes the text usable for a study like the present.[9]

Let us return to the testing of the influence of centering status on the choice of referring expression. The search for DemPr showed no clear correlation between the choice of referring expression and centering status. In (4) there are some typical examples for sentences containing DemPr.

---

9. On the value of Otfrid as an object of linguistic investigation see e.g. Somers Wicka (2009).

(4) a. (context: "Iuo búah", quad, "wéizent \ thaz mán ouh góta heizent; \\ giwisso ságen ih iz iu \ thaz man sie nénnit thar zi thíu – "your books", he said, "tell you that men are also called gods; Verily, I tell you, that one calls them so for the following reason.")

A: *Nu   thie zi góte sint ginánt, thie búent hiar*
     Now who to God are called   these farm here
     this wóroltlant,
     this earth

B: <u>*then*</u>   *gótes wort gizáltun waz sie iu io*
     to those God's words told   what they you ever
     ságen scoltun
     say   should

'Well, they that are called after God, they farm this earth. God's words tell them, whatever they should say to you.'   (Otfrid, Ev., 3,22,49–52)

Modern German version:

Die, die nach Gott genannt sind, die bebauen die Erde. <u>*Denen*</u> *weist Gottes Worte, was sie euch sagen sollen.*

b. A: *thie hohun áltfatera   éntont anan kúninga*
       the high   patriarchs end   with king

B: *thiu thrítta zuahta   thánana   tház warun édilthegana*
     the   third   generation from-there that were   noblemen

C: <u>*Thie*</u> *warun wúrzelun thera   sáligun blúomun*
     these were   roots   of-the blessed flower

'The honourable patriarchs end when a king arises. The third generation after that was noble. They were the roots of the blessed flower (= Mary).'
                           (Otfrid, Ev., 1,3,25ff.)

Modern German version:

Die Zeit der hohen Patriarchen endet mit einem König. Die dritte Generation danach waren Edelleute. <u>*Die*</u> *waren die Wurzel der gesegneten Blume.*

c. (context: Then námon er irkánta \ só man nan ginánta – he recognised the name as soon as one said it)

A: *tho gab er zo ántwurte tház, thaz ér ther selbo mán*
     then gave he to answer   that   that he the   same man
     ni   was
     not was

B: *'her gómo then   ir   záltut   joh námahafto nántut ni*
     The man   whom you pursued and by-name   called not
     bin ih <u>thér;</u>
     am I   DemPr

'Then he (= John the Baptist) answered that he was not this man. "The man whom you looked for and whom you called by name, I am not him."'

Modern German version:

Dann gab er das zur Antwort, dass er dieser Mann nicht sei. "Der Mann, den ihr sucht und mit Namen ruft, <u>der</u> bin ich nicht.

<div align="right">(Otfrid, Ev., 1,27,25–28)</div>

d. A: *Nichódemus ther gúato, er quám thar tho gimúato,   unz*
       Nicodemus the   good   he came there then pleasantly until
       ér nan tho   thána nam, ther náhtes er   ju   zi
       he him then thence took who at-night before already to
       ímo   quam;
       him   came

    B: *<u>Ther</u> bráng   mit imo in wára sálbun   filu   díura*
       He   brought with him indeed ointments many expensive

<div align="right">(Otfrid, Ev., 4,35,17ff.)</div>

'Then the good Nicodemus came there with pleasure, when he had taken him (= Jesus) from it (= the cross). He came to him already at night. It is true, he brought with him many expensive ointments.'

Modern German version:

Dann kam der gute Nikodemus mit freundlichem Sinn, bis er ihn da ab-nahm, der bereits in der Nacht zuvor zu ihm kam. #<u>Der</u> brachte mit sich tatsächlich sehr teure Salben.

e. A: *Thaz mári ward ouh   mánagfalt ubar Júdeono lant, ubar*
       that fame was   also numerous over Jews'   land over
       líuti   manage, thie fúarun   al zisámane,
       people many   these travelled all together

    B: *Sie   gérotun al bi mánne   inan zi rínanne, […] joh*
       They desired all man-by-man him to touch   also
       sih   zen sínen guatin io   étheswaz   gifuagtin.
       itself to   their welfare sometime something fit._{subj.3.PL.}

    C: *<u>Thie</u>   bráhtun imo ingégini síechero   manno   ménigi,*
       DemPr brought him towards sick._{GEN.PL.} men._{GEN.PL.} multitude

<div align="right">(Otfrid, Ev. 2,15,5–9)</div>

'That fame went through all Judea, to all people; they travelled all [to him]. They desired all, man by man, to touch him, so that at some time some-thing might arise from it for their recovery. They brought towards him a lot of sick people.'

In (4a), a DemPr is used in a Retain relation. The referent of *then* (the group of people farming the soil) are the Cb of the sentence B, but they are not in subject function, so that they are not the Cp of the sentence B, which would be *gótes wort*. They represent, however, the Cb of the preceding sentence A, as the same group of people referred to in A is present in the verse before A. In A it is referred to with

some additional predications in form of a Left Dislocation (which is the reason that the DemPr may be used in A, although they have been referred in the preceding verse via *sie*). This usage of the DemPr as indicator of a Retain relation is quite typical and it is felicitous also in Modern German, as can be seen from the modern German version given under the example. It is also hinted at e.g. in the result from Bosch et al. (2003) and Bosch et al. (2007), that PerPr highly prefer nominative antecedents while DemPr tend to have non-nominative antecedents.

Likewise, the usage demonstrated in (4b) is also felicitous in Modern German. Here the DemPr is correlated with a Smooth Shift relation: The Cb of C, the noblemen of the third generation, is at the same time the Cp of C. It is not coreferent with the Cb of sentence B, the patriarchs of sentence A, which are hidden in *thanana* 'thereafter' in sentence B.

Demonstrative pronouns could also be used under certain circumstances in Continue relations, namely if there is a contrastive context. This usage is also felicitous in Modern German. An OHG example is presented in (4c). The DemPr in B is subject, thus Cp, and is also referring to a referent in A that is present in the preceding discourse. It is questionable whether *ther selbo man* 'the same man' counts as Cb in A, though. So possibly here is a Smooth Shift relation. Anyway, we can say that there is some local disturbance in the thematic structure, which is mostly brought about by the continuous contrastive context where the topic shifts from John the Baptist (who is the man giving the answer) and Jesus (who is the man John is asked about and who he says he is not). Note that this is not a case of Left Dislocation, in which case we would expect the usage of a DemPr anyway, as the resumptive pronoun is not in the prefield.[10]

Example (4d), however, demonstrates a usage of DemPr that is so not felicitous in Modern German. The referent of the DemPr, Nicodemus, is Cb and Cp of sentence B. He is not the Cb of the preceding utterance A, however, as Nicodemus is newly introduced in this utterance. So again it looks as if there is a Smooth Shift relation. Note however that Nicodemus gained such salience that in Modern German he can only be referred to by a PerPr and not a DemPr. As is well known (s. e.g. Abraham 2002: 461; Winter 2003: 33f.), in an anaphoric chain of the form $a(1)_i, a(2)_i, \ldots, a(n)_i$ the following holds: if a PerPr is used as anaphoric expression in $a(m)_i$, where $1 \leq m \leq n$, no DemPr can be used for $a(m + p)_i$, where $1 \leq p \leq n$. In other words: Once a personal pronoun is used for referring to some entity, the same entity cannot be referred to by a DemPr. This follows directly from Ariel's (1990) accessibility scale. Examples like (4d) are a direct violation of this constraint. The

---

10. On the distinction of Left Dislocation and Hanging Topic see e.g. Altmann 1981; Shaer & Frey 2004)

fact that such examples occur hint at the possibility that salience is possibly not an exclusive factor relevant for anaphor resolution in OHG. That this 'thematic' usage of the DemPr is not idiosyncratic to this special line in the text is shown e.g. by similar examples in Tatian that are discussed in Abraham (2007b, 251f.).

Whereas in (4d) it is still possible to interpret the relation as a shift, there are clear examples of Continue relations with a DemPr. Consider (6) below. Here the group of people in A (*liuti manage*) is Cb in B, referred to by PerPr, and also Cp, as it is the subject. Nevertheless, in C the same group is referred to by *thie*, although they continue to be both Cb and Cp. The usage of a DemPr in the Modern Standard German version is infelicitous, as can be seen from the translation, although in many dialects the use of demonstrative pronouns in such cases is fine.[11]

So the search for DemPr found examples for DemPr in all Centering relations. A problem of this method is, of course, that many cases of DemPr are ambiguous. The d-pronoun *ther, thie, thaz* can be used both as simple demonstrative pronoun and as the relative pronoun. We saw this ambiguity in some examples under (4), e.g. in the C clause of (4b), which could be either a relative clause with verb second syntax or an independent clause introduced by a demonstrative pronoun. In general this ambiguity is however not problematic. As long as the clauses starting with a demonstrative or relative pronoun are not clearly embedded, they have to be regarded as independent propositions between which coherence relations hold. That an analysis of clauses which start with a d-pronoun as DemPr is at least possible show examples like (5) from the Monsee Fragments.

(5) *Gotes minni ist gagozan in unsere muotuuilun durah heilagan geist*
God's love is shed in our own-will by holy spirit
der uns gageban uuarth; Huuanta ano <u>*dea*</u> *nist dir*
who us given became For without DemPr not-is to-you
eo uuiht bidarbi des du hapen maht
anything useful of-which you have may
'God's love is shed in our heart by the Holy Spirit, who was given to us. For without it (= God's love) nothing that you have is of any use to you.'
(Monsee Fragments MF,2_VG.XXIX,10)

In (5) the d-pronoun *dea* is clearly a DemPr. The clause cannot be a relative clause as it is introduced by the conjunction *huuanta* 'because, for'. This is, by the way, an example for a Retain relation.

---

11. The paradigm of the personal pronoun in non-clitic usage has adopted forms of the demonstrative pronoun in these dialects, e.g. Zurich German (cf. Nübling 1992: 265). The distribution of personal pronoun and demonstrative pronoun follows other factors in such cases.

Let us turn to the Centering analysis of selected passages of Otfrid. Here the impression we already got from the search for DemPr was confirmed, namely that there is no direct correlation between centering relation and choice of anaphoric expression. Table 3 shows the rate of different kinds of referring expressions for different coherence relations.

Table 3. Rate of referring expressions for coherence relations in selected passages from Otfrid

	Continue		Retain		Smooth Shift		Rough Shift	
	n	*% exp./rel.*	n	*% exp./rel.*	n	*% exp./rel.*	n	*% exp./rel.*
zero pronoun	5	*9.4*	0	*0*	2	*7.1*	0	*0*
PerPr	42	*79.2*	10	*90.9*	13	*46.4*	2	*40.0*
DemPr	1	*1.9*	0	*0*	0	*0*	2	*40.0*
full NP	5	*9.4*	1	*9.1*	13	*46.4*	1	*20.0*
sum	**53**	*100*	**11**	*100*	**28**	*100*	**5**	*100*

There is no clear correlation, but at least some tendencies. The relations Continue and Retain, in which the Cb of a discourse segment is identical to the Cb of the preceding segment, show a high rate of PerPr, whereas the rate in the Shift relations is below 50%.

Table 4 shows the rate of different coherence relations for different kinds of referring expressions.

Table 4. Rate of coherence relations for referring expressions in selected passages from Otfrid

		Continue	Retain	Smooth Shift	Rough Shift	sum
		n	n	n	n	
zero pronoun	n	5	0	2	0	7
	% rel./exp.	*71.4*	*0*	*28.6*	*0*	*100*
PerPr	n	42	10	13	2	65
	% rel./exp.	*64.6*	*15.4*	*20.0*	*3.1*	*100*
DemPr	n	1	0	0	2	3
	% rel./exp.	*33.3*	*0*	*0*	*66.7*	*100*
full NP	n	5	1	13	1	20
	% rel./exp.	*25.0*	*5.0*	*65.0*	*5.0*	*100*

Here the same picture prevails: No clear correlation, but clear tendencies. With zero pronouns and PerPr, the rate of continue relations is rather high. In both cases the second highest rate are shift relations. Whereas in the case of zero pronouns this is not surprising (zero pronouns are nearly exclusively in subject function in OHG), in the case of PerPr it is worth noting. As Continue and Smooth Shift are the relations in which the Cb is at the same time the Cp, there is some correlation to subjecthood visible, but not in the sense of the Subject Generalization from above, but rather such that resumption in subject function is correlated to PerPr, independent of the function of the antecedent. DemPr, rare as they are, show in the data sample no clear correlation to coherence relations: 2 of them are associated with rough shift, but one with continue.

To get an idea what little association there is between coherence relations and the choice of referring expression, consider Example (6), a continuous text sample from the analysed parts of Otfrid.

(6) A: *Thaz mári   ward ouh mánagfalt   ubar Júdeono lant,   ubar*
       the   tidings was also multifarious over Jews'   country over
       líuti   manage, thie fúarun al zisámane
       people many   these went   all together
    B: *Sie   gérotun al bi mánne inan zi rínanne joh Ø sih*
       They desired all by man   him to meet   and   themselves
       zen sínen guatin   io   étheswaz   gifúagtin.
       to   their goodness ever something link
    C: *Thie bráhtun imo ingégini síechero manno ménigi,   bifangan*
       these brought him towards sick   men's   multitude captured
       mit únmahtin   joh míssilichen súhtin;   (Otfrid, Ev. 2,14,5–10)
       by   weaknesses and diverse   sicknesses

    'The tidings (= that Jesus was there) became widely known throughout Judea and by many people. They went all together (sc. to see him). They all desired to meet him, one by one, and wanted to add something to their own recovery. They brought to him a lot of sick people, with all sorts of ailments and weaknesses.'

    Modern German version:

    Sein Ruhm verbreitete sich in ganz Judäa, zu vielen Leuten, die alle zu ihm hinfuhren. Sie wollten alle, jeder einzelne, ihn berühren, damit sie auch irgendwann etwas zu ihrer Genesung hinzufügen mögen. #Die brachten ihm eine Menge Kranker entgegen, belastet mit Schwächen und allerlei Krankheiten.

We see here a wild mixture of DemPr, PerPr and zero pronoun for reference to the same entity, the group of people introduced in the second line (*liuti manage*) which is the Cb in all cases. The first pronominal reference in the last clause of A need not be exceptional, as this might well be an example of a relative clause, demonstrative pronouns being used as relative pronouns also in OHG. But the use of the demonstrative in C, which looks like a verb-second clause (the pronominal *imo* after the verb shows that the clause order cannot be the result of extraposition) and thus is probably not a relative clause, remains unexpected, after the same entity has been referred to by personal pronouns, even zero pronouns before.

Examples like (6) show that the reference by PerPr is quite independent from centerhood. The only referent that is clearly associated to PerPr is Jesus (referent of *inan* in the 3rd line and *imo* in the 5th line), although he is not the Cb. Jesus is not even mentioned in the A sentence. This is interesting for the following reason: Following Abraham (2002) for Modern German and Petrova and Solf (2010) for the OHG of the Tatian, one might assume that the choice of anaphoric expressions is associated with topichood in a non-centering sense (that is, the topic, not the Cb). But Example (6) illustrates that topichood is not a sufficient condition on the choice of referring expression either: Jesus is not the topic of the passage, at least not on a local level (he might count as discourse topic, though). This shows that the choice of anaphoric expressions in OHG depends on quite different factors than it does in Modern German, among which are possibly animacy and an urge to treat *nomina sacra* different from 'normal' referents.[12]

We can use still another test. Winter (2003) posits for modern German clear rules: If it is ambiguous to which antecedent a Cb is coreferent, a PerPr selects the antecedent that leads to a more coherent relation, whereas a DemPr selects the antecedent that leads to a less coherent relation. In her corpus study, she gets a success rate of 87.76% – that is: 87.76% of the cases in her corpus confirm to her rules among which the one mentioned above is the central one (Winter 2003: 44). I looked how the success rate was in the OHG text passages. The rate was almost identical: 34 out of 39 Cbs in which the reference was ambiguous conformed to Winter's rules, which is a rate of 87.18%. So the tendency is equally strong as in Modern German, but there are stark differences in what directions deviations from this rule occur. In (6) we saw an example of DemPr for contexts in which we would expect a PerPr nowadays. There are also examples for PerPr in which we would not expect a PerPr today (7).

---

12. This idea is elaborated in Speyer (2016).

(7) a. *Gibót   si   then   sar gáhun then   thes   lídes   sahun so*
Ordered she those at-once   those that.$_{gen}$ misery.$_{gen}$ saw   so
wás   so   er in   giquati iz íagiliher dati.   Thar stuantun
what so he them told.$_{subj}$ it everyone did.$_{subj}$ there stood
wázarfaz, […] warun stéininu thiu fáz   siu   mohtun wéren
water-jugs   were of-stone the jugs they could   last
thes   thiu   báz. Gibót   tho selbo drúhtin, siu   wátares
the.$_{gen}$ the.$_{instr}$ better ordered then same Lord   they water.$_{gen}$
irfúltin;   thaz dátun sie   giwurtig unz in óbanentig.
filled-in.$_{subj}$ that did   they gladly till in rim
'She (= Mary) at once ordered those who saw the misfortune to do what-
ever he would tell them. There stood some jugs of water … the jugs were
made from stone, so that they could last longer. The same Lord ordered
them that they filled water into [sc. the jugs]. They did this gladly, up to
the rim.'   (Otfrid, Ev. 2.8.25–27a, 34–36)

b. *Séhs dagon fora   thíu quám er zi Bethániu, thar er fon*
six   days before this came he to Bethany   there he from
tóthe irwácta,   Lázarum irquícta.   Tho zemo ábande
death awakened Lazarus   made-alive Then to-the evening
sár   gáretun   sie   sin múas thar;
quickly prepared they his meal there
'He came to Bethany six days before that, where he had awakened Lazarus
from the death. Then, on that evening, they prepared quickly a meal for
him   (Otfrid, Ev. 4.2.5–7)

In (7a) the referent of *siu* in l.35 are the servants mentioned in l.25f. The whole
passage in between deals with the water jugs for washing the feet. Yet in l.35 the
servants are referred to by a PerPr and acquire Cb-status in l.36. The relation is a
Smooth Shift relation, as the Cb of l.35 is the jugs. We would expect the servants to
be re-introduced by a full NP in l.35.

There is however a way to explain the usage of PerPr in this context. The passage
l.27–34 is an elaboration on the water jugs; basically a self-contained sub-discourse.
If we assume that centers look for the preceding discourse segment on the same
hierarchical level (cf. Grosz & Sidner 1986; Speyer 2007), l.26 and l.35 are adjacent
on the same hierarchical level. If that is so, the relation is not a Shift relation, but
a Continue relation, and the usage of a PerPr is according to Winter's rule. This
shows that the hierarchical discourse organization plays a larger role for anaphora
resolution in OHG than it does in Modern German.

Whereas in (7a) it is possible to explain the usage of PerPr, there are some ex-
amples of quirky reference that simply cannot be accounted for. Take (7b) for exam-
ple. The referent of *sie* in l.7 is the family in Bethany, viz. Lazarus, Mary and Martha.

Yet only Lazarus has been mentioned in the passage. The group as a whole is only inferable. Here reference by a PerPr would not be felicitous in Modern German.

The variation in the usage of DemPr and PerPr in OHG can be interpreted such that the rules for the division of labor between the pronouns have not yet been settled. As the PerPr (having been a weak demonstrative before) only acquired its role as PerPr relatively shortly before, this is not really surprising.[13]

In sum, we can say that there is no direct correlation between coherence relations in Centering terms and the choice of anaphoric expression, although there are clear tendencies. Continue relations are mostly associated with PerPr, whereas DemPr are seldom associated with Continue, but they can be. Note that this result is not unexpected: Schlachter (2012) shows for her corpus (Isidor and Monsee fragments) that zero pronouns are compatible with all coherence relations, even shifts. Note that zero pronouns should be even more suitable for marking 'high' coherence, that is: continue, if we believe accessibility scales in the style of Ariel (2001). They work quite well for Modern German (see Ahrenholz 2007: 243, Ellert 2010), so they should work for OHG as well, especially considering the fact that they are regarded as a language universal. Combining the results of this study with Schlachter (2012) we can be quite confident in stating that centering has no clear effect on the choice of anaphoric expressions. If an accessibility scale works for OHG, the factors determining the accessibility (or the salience) are probably not centerhood and probably not even topichood, but some other factors which might in the end show some correlation to centerhood, however, thus explaining the visible tendencies.

4.   Later periods

The modern system is more or less in operation in Early New High German (ENHG). I looked at two texts from that period, namely the *Buch der heiligen Altväter* and Hans Mair's *Troja*. Both texts are from the early ENHG era, between 1350 and 1400, and both texts are from the Swabian dialect area (the *Buch der heiligen Altväter* from Reute monastery near Ravensburg, Mair's *Troja* from Nördlingen).

Both texts show a similar treatment of DemPr and PerPr: DemPr are associated with Shift relations, PerPr are associated with Continue. This mirrors the modern state of affairs as outlined in Section 2. Some traits remind of the OHG state of affairs, however, for instance that the hierarchical discourse organization plays a larger role than today. Consider (8). The hierarchical structure of (8) is given below the example.

---

13.   A similar conclusion draws Abraham (2002).

(8) A: *in ainem iaͬ͘ do gieng er (= Hilarion) vs nach siner gewonhait*
    in one year then went he   out after his custom

    B: *vnd kam gen ainer statt ciluza genant*
    and came to a city Ciluza called

    C: *vnd was haidesch*
    and was gentile

    D: *in der was ainer hand haiden*
    in this was one hand heathen

    E: *die haissent sarraceni*
    they are-called Sarracenes

    F: *den hett er och vil gůtes getaͬ͘n*
    to-these had he also much good done

                                (Buch Altväter 67b,23–68a,4)

'One year he traveled, as he was used to, and came to a city called Ciluza. This was a gentile city. One kind of heathen lived in it. They were called Sarracenes. He had done much good to them before.'

Modern German version:

In einem Jahr reiste er umher, nach seiner Gewohnheit, und kam in eine Stadt namens Ciluza. Die war heidnisch, in der waren eine bestimmte Sorte Heiden, die heißen Sarrazenen. Denen hatte #er / √Hilarion viel Gutes getan.

Hierarchical structure: A → B   (→)   F

(diagram: A → B, with arrow down from B to C; C → D → E; arrow up from E to F)

In the F sentence there are two referents that are realised anaphorically, Hilarion and the Sarracenes. The F sentence is linearly adjacent to E, which is part of a subdiscourse comprising sentences C to E. It is, however, adjacent on the same hierarchical level to sentence B. In relation to sentence B, a Continue relation would hold with Hilarion as Cb, as Hilarion is the Cb also of B and is the subject of F. In relation to the linearly adjacent sentence E, however, a Retain relation holds, as the Sarracenes are the Cb of E, but they are not the Cp of F. In accordance to Winter's rules, the PerPr is chosen for reference to the potential Cb that stands in the most coherent relationship to the preceding discourse segment, i.e. Hilarion, and the DemPr is chosen for reference to the other potential Cb that stands in a less coherent relationship, i.e. the Sarracenes. This consideration works however only as long as the non-linear relationship between B and F counts for Centering reasons. In Modern German, the role of the hierarchical organization is still there, but not as strong as in ENHG, as can be sen from the Modern German version of the passage. Here reference by a PerPr to the Cb of the hierarchically adjacent discourse segment is somewhat possible, but dispreferred; more felicitous is reference by a full NP in such cases.

An important point that links ENHG to Modern German rather than to OHG is that DemPr are not associated with Continue any more. This is also shown by (8). Related to that is the fact that Winter's rule is in operation as it is in Modern German. An example of a linearly progressing discourse, (9), illustrates this point.

Here we have two referents, Jason and the ram. The B sentence presents both referents, one realised as PerPr, one realised as DemPr. Both referents were mentioned in the A sentence. If Jason, the referent of the PerPr, would be the Cb, a Continue relation would hold between A and B. If the ram, the referent of the DemPr, would be the Cb, a Smooth Shift relation would hold, as the ram is newly introduced in A and thus cannot be the Cb of A. In Centering Theory, the ranking of coherence relations introduced at the end of Section 2 has also the consequence that the higher ranked relations are preferred over the lower ranked, as they involve less computational effort (Walker et al. 1998: 5ff.; Winter 2003: 13).[14] So the Continue relation would be preferred in this case; the PerPr indicates the preferred relation, just according to Winter's rule.

(9) A: *do   er (= Jason) daz allez hät überwunden, do   gieng er*
        when he   that all   has surmounted then went   he
        frölich und mänlich   zu dem wider, an dem   er kain
        happy   and masculine to the   ram   at which   he no
        wer   nit envand,
        resistance not NEG-found.
     B: *den   nam er bi den hornen,*
        DemPr.ACC took he by the horns
     C: *und stach   in   ze tod,*
        and stabbed him   to death
     D: *und zoch im   ab   sein guldin gewand,*
        and drew him   away his golden garment

                                                    (Mair: Troja, 25,10–13)

Note that the referent of the DemPr can be promoted to PerPr, if the original referent of the PerPr is promoted to zero, as is the case in the C and D sentences of (9).

A little detail should be mentioned here. Note that in (9) the more cohesive anaphoric expression (PerPr in A, B, zero in C, D) is always in subject function, whereas the less cohesive anaphoric expression (DemPr in B, PerPr in C, D) is always in a function different from the subject. In fact, the ENHG examples I looked at all followed this generalization. This can be interpreted as a hint to an answer to the notorious question what factor is relevant for the ranking of Cfs. If the Cfs are ranked according to their syntactic function, as Walker et al. (1998) propose

---

14. This is in parallel with more general assumptions, e.g., that discourses come with the presumption of being coherent (Bublitz, Lenk & Ventola 1999).

for English, this observation would follow directly. Another fact that points in the same direction is that we should get examples in which the most cohesive pronoun has another syntactic function, as nothing would prevent this, if the ranking of Cfs were not determined by the syntactic function. In Modern German, the picture seems to be less clear than in ENHG (or at least these two ENHG texts), but there is a strong tendency.

5.   Conclusions

The question whether anaphora resolution in OHG was sensitive to centerhood in general cannot be answered totally affirmative. Although the centering analysis of Otfrid's *Evangelienbuch* revealed strong tendencies, e.g. that Cbs in Continue relations tend to be realised by PerPr, whereas Cbs in Retain and Shift relations tend to be realised by DemPr, the association is far from categorical. The strong tendencies are most likely due to some factor determining salience that is mediately connected to centerhood, thus leading to a relatively good match of anaphoric expressions to coherence relations. That the system is less sensitive to centerhood than in Modern German is indicated by the fact that there are blatant violations of generalizations that hold for Modern German, such as Winter's rule (in cases of ambiguous reference, PerPr signal a more coherent relation, DemPr a less coherent one) or the rule that in an anaphoric chain no DemPr may occur once a PerPr has been used. The modern system that is somewhat sensitive to centerhood is in operation at latest since the 14th century, as is evidenced by centering analyses of two texts from this period. As a matter of fact, the system seems to have been more sensitive to centerhood than today, defining locality not purely linearly but in dependency of hierarchical discourse organization, and using grammatical function as relevant factor for the Cf ranking.

References

Abraham, W. 2002. Pronomina im Diskurs: Deutsche Personal- und Demonstrativpronomina unter Zentrierungsperspektive. Grammatische Überlegungen zu einer Teiltheorie der Textkohärenz. *Sprachwissenschaft* 27: 447–491.

Abraham, W. 2007a. Discourse binding. DPs and pronouns in German, Dutch, and English. In *Nominal Determination. Typology, Context Constraints, and Historical Emergence* [Studies in Language Companion Series 89], E. Stark, E. Leiss & W. Abraham (eds), 21–47. Amsterdam: John Benjamins.

Abraham, W. 2007b. The discourse-functional crystallization of the historically original demonstrative. In *Nominal Determination. Typology, Context Constraints, and Historical Emergence* [Studies in Language Companion Series 89], 241–256. Amsterdam: John Benjamins.

Ahrenholz, B. 2007. *Verweise mit Demonstrativa im gesprochenen Deutsch. Grammatik, Zweitspracherwerb und Deutsch als Fremdsprache*. Berlin: De Gruyter.

Altmann, H. 1981. *Formen der Herausstellung im Deutschen*. Tübingen: Niemeyer.

Ariel, M. 1990. *Accessing NP antecedents*. London: Routledge.

Ariel, M. 2001. Accessibility theory: An overview. In *Text representation*, T. Sanders, J. Schliperoord & W. Spooren (eds), 29–87. Amsterdam: Benjamins.

Axel, K. 2007. *Studies on Old High German Syntax. Left Sentence Periphery, Verb Placement and Verb-second* [Linguistik Aktuell/Linguistics Today 112]. Amsterdam: John Benjamins.

Axel, K. & Weiß, H. 2011. Pro-drop in the history of German. From Old High German to the modern dialects. In *Null Pronouns*, P. Gallmann & M. Wratil (eds), 21–51. Berlin: De Gruyter.

Bethke, I. 1990. *Der, die, das als Pronomen*. München: Judicium.

Bosch, P. 2012. Parameters of anaphoric reference for German personal and demonstrative pronouns. Talk given at 34th DGfS Annual Meeting, Frankfurt, March 8, 2012.

Bosch, P., Rozario, T. & Zhao, Y. 2003. Demonstrative pronouns and personal pronouns. German der vs. er. Proceedings of the EACL 2003, Budapest.

Bosch, P., Katz, G. & Umbach, C. 2007. The non-subject bias of German demonstrative pronouns. In *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference* [Studies in Language Companion Series 86], M. Schwarz-Friesel, M. Consten & M. Knees (eds), 145–164. Amsterdam: John Benjamins.

Bosch, P. & Umbach C. 2007. Reference determination for demonstrative pronouns. *ZAS Papers in Linguistics* 48: 39–51.

Bublitz, W., Lenk, U. & Ventola, E. (eds). 1999. *Coherence in Spoken and Written Discourse* [Pragmatics & Beyond New Series 63]. Amsterdam: John Benjamins.

Cardinaletti, A. & Starke, M. 1996. Deficient pronouns: A view from Germanic. In *Studies in Comparative Germanic Syntax, Vol. II*, H. Thráinsson, S. D. Epstein & S. Peter (eds), 21–65. Dordrecht: Kluwer.

Chiarcos, C., Claus, B. & Grabski, M. 2011. Introduction: Salience in linguistics and beyond. In *Salience: Multidisciplinary Perspectives on its Functions in Discourse*, C. Chiarcos, B. Claus & M. Grabski (eds), 1–28. Berlin: De Gruyter.

Eggenberger, J. 1961. *Das Subjektpronomen im Althochdeutschen. Ein syntaktischer Beitrag zur Frühgeschichte des deutschen Schrifttums*. Grabs: Selbstverlag.

Ellert, M. 2010. *Ambiguous Pronoun Resolution in L1 and L2 German and Dutch*. PhD dissertation, MPI Nijmegen.

Grosz, B. & Sidner, C. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12: 175–204.

Grosz, B., Joshi, A. K. & Weinstein, S. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 21: 203–225.

Gundel, J. K., Hedberg, N. & Zacharski R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274–307.

Nübling, D. 1992. *Klitika im Deutschen. Schriftsprache, Umgangssprache, alemannische Dialekte*. Tübingen: Narr.

Petrova, S. & Solf, M. 2010. Pronominale Wiederaufnahme im ältesten Deutsch. In *Historische Textgrammatik und historische Syntax des Deutschen*, A. Ziegler & C. Braun (eds), 339–363. Berlin: De Gruyter.

Primus, B. 2011. Animacy, generalized semantic roles, and different object marking. In *Case, Word Order, and Prominence. Interacting Cues in Language Production and Comprehension*, M. Lamers & P. de Swart (eds), 65–90. Dordrecht: Springer.

Schlachter, E. 2012. *Syntax und Informationsstruktur im Althochdeutschen. Untersuchungen am Beispiel der Isidorgruppe*. Heidelberg: Winter.

Shaer, B. & Frey, W. 2004. 'Integrated' and 'non-integrated' left-peripheral elements in German and English. *ZAS Papers in Linguistics* 35: 465–502.

Somers Wicka, K. 2009. *From Phonology to Syntax: Pronominal Cliticization in Otfrid's Evangelienbuch*. Tübingen: Niemeyer.

Speyer, A. 2007. Die Bedeutung der Centering Theory für Fragen der Vorfeldbesetzung im Deutschen. *Zeitschrift für Sprachwissenschaft* 26: 83–115.

Speyer, A. 2016. Factors determining the choice of anaphora in Old High German – A survey of zero and personal pronoun usage in Otfrid. In *Experimental Perspectives on Anaphora Resolution*, A. Holler & K. Suckow (eds), 113–141. Berlin: De Gruyter.

Stark, E., Leiss, E. & Abraham, W. (eds) 2007. *Nominal Determination. Typology, Context Constraints, and Historical Emergence* [Studies in Language Companion Series 89]. Amsterdam: John Benjamins.

Strube, M. & Hahn, U. 1996. Functional centering. *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, 270–277.

Turan, Ü.D. 1998. Ranking forward-looking centers in Turkish: Universals and language-specific properties. In *Centering Theory in Discourse*, M. A. Walker, A. K. Joshi & E. F. Prince (eds), 139–160. Oxford: OUP.

van Dyke, J. 2012. The role of memory in language and communication. In *Cognition and Acquired Language Disorders: A Process-oriented Approach*, R. Peach & L. Shapiro (eds), 94–120. St. Louis: Mosby Elsevier.

Walker, M. A., Joshi, A. K. & Prince, E. F. 1998. Centering in naturally occurring discourse: An overview. In *Centering Theory in Discourse*, M. A. Walker, A. K. Joshi & E. F. Prince (eds), 1–28. Oxford: OUP.

Winter, Y. 2003. Parameter zur Bestimmung der Referenz von Demonstrativpronomina in Texten. Versuch einer Analyse mit Hilfe der Centering Theory. BA thesis, Osnabrück.

On the processing of Free Indirect Discourse

First results and methodological challenges

Susanna Salem, Thomas Weskott and Anke Holler
University of Göttingen

In this contribution we report on the results from two psycholinguistic experiments investigating the processing of Free Indirect Discourse (FID). We conceive of FID as a linguistic means that cues comprehenders to take over the perspective of a protagonist in third-person narrations. Using both on-line and off-line measures, we tested the hypothesis that the referent of the protagonist receives a higher activation status during reading if his or her thoughts are related through FID. The FID cues we used were questions and discourse particles. In addition, we compared different inferential statistic procedures in the analysis of the results. Although the cues that were employed as FID markers in the experimental materials had an influence on the perception of narrative perspective, no indication was found for the hypothesis that narrative perspective mediated through FID influences the salience of the protagonist during reading. We discuss the implications of this null result and point to some more general methodological problems arising in the investigation of processing of literary text.

1. Introduction

There are numerous means that a speaker/author might use to perspectivize the events described in a narrative. A prominent one is Free Indirect Discourse (FID), which is used to report on a protagonist's attitudes towards some situation or event. As the comparison of the cases in (1) shows, FID (1c) shares features of both Direct (1a) and Indirect Discourse (1b):

(1) a. He woke up and thought: "What the hell am I doing here?"
    b. He woke up and asked himself what (*the hell) he was doing here.
    c. He woke up. What the hell was he doing here?

While the FID version in (1c) shares the shifted tense and pronoun with the indirect report in (1b), it also differs from (1b) with respect to its potential to retain the

expressive material ('the hell') from the original thought. This mixed nature of FID, and how it can be accounted for by semantic theories, has recently received a lot of attention, reviving the pioneering work of Banfield (1973, 1982); see e.g. Schlenker (2004), Sharvit (2008), Maier (2012), and Eckardt (2014). This line of research has generated a number of interesting hypotheses about the semantic representations of FID, and, consequently, about its on-line comprehension.

While there have been attempts to investigate FID experimentally (mostly from a narratological perspective, with the notable exception of Kaiser and Cohen 2012, and Harris 2012, 2014), our understanding of the psycholinguistics of FID, that is, its processing and representation, is still lagging behind the formal approaches. This has only partly to do with methodological problems earlier studies were beset with, but also with the fact that the effects of FID on processing are, as we are about to argue, rather subtle.

This contribution has two aims: firstly, to present two psycholinguistic experiments that were conducted to validate a hypothesis about the processing of FID. While FID has been extensively studied by literary scientists at least since Genette (1994; see also Fludernik 1993), it has lately attracted the interest of formal semanticists (cf. the references above), psychologists (Bortolussi & Dixon 2003), and psycholinguists (Kaiser & Cohen 2012; Harris 2012, 2014). In the two experiments reported here, we investigated whether FID influences the activation status of the entry for the protagonist in memory during reading. We hypothesized that FID would make the protagonist referent more accessible, given that in a passage containing FID, the situation is described from his or her perspective. This was expected to result in a processing advantage every time the protagonist referent has to be retrieved from memory, for example when a coreferential pronoun has to be resolved, or when answering questions about the protagonist. Being largely compatible with results reported in the literature, our findings indicate that, while off-line effects of FID seem to be quite robust, pinning down the influence of FID in experiments using on-line measures seems to be rather difficult.

Our second aim, if only a modest one, is to take the opportunity of the two experiments we conducted to perform a methodological exercise. We hope to be able to show that more recent developments in inferential statistics can help researchers of narratives (and, possibly, beyond) to avoid spurious generalizations. We take this to be an important point to make especially in situations where rather subtle effects may be blurred by large variances, as is often the case when studying the processing of larger stretches of text. But, to make an obvious caveat right up front: we do not think that the kind of spurious generalization we are thinking of here is specific to experimental research on narratives. (In fact, we think that the problems we point to are rather pervasive throughout the whole of experimental

psycholinguistics.) Given that experimental studies of text comprehension necessarily have to deal with more variance than, say, studies of sentence processing, and that effects of pragmatic phenomena like FID can be quite subtle, it seems reasonable to use all the statistic tools available to partition and explain the variance present in the data. More specifically, we will show how the use of advanced statistical tools affords us with insights into the brittleness of on-line FID effects in the face of large between-item variance.

Moreover, we want to address in this contribution the recent criticism concerning the "researcher degrees of freedom" in data treatment and analysis brought forward by Simmons, Nelson, and Simonsohn (2011). They argue that the danger of getting false positive results (Type I errors) increases with the number of decisions between, for example, different statistic procedures. In the worst case, the authors show, the maximum probability of producing a false positive result (usually set to 5%) may increase to up to 60%, depending on the degrees of freedom. These degrees of freedom pertain to decisions like reporting raw vs. transformed reaction times; outlier removal vs. no outlier removal; ANOVA vs. linear mixed effects models, and so on. By reporting different analyses for our data sets, we want to make these decisions transparent, and show how they influence the significance of our results. By this, we hope to contribute to a better understanding of the role of statistic procedures in experimental research on narratives, but also beyond.

2.   Theoretical background

While FID has been given many different labels in various languages and research contexts (*discours indirect libre, erlebte Rede*, to name only the French and German terms), and has been studied under numerous headings such as *perspectivization, dual voice, point of view, stream of consciousness, focalization*, etc., what seems to lie at its heart is that it is a linguistic device by means of which a third-person narrator can report the mental contents (mental attitudes like beliefs, desires, etc.) of a protagonist. The reported content is said to be reported indirectly, because FID lacks an explicit factive component like "Peter said …", or "…, Jane wondered". Semanticists have been concerned with the semantics of FID, trying to account for the problems originating from its being a hybrid between direct and indirect speech/thought, from both of which it borrows grammatical features. From a psycholinguistic point of view, it is an interesting phenomenon because it exhibits a peculiar relation between a rather innocent surface cue for interpretation, and a considerable interpretive effect this cue exerts for the ultimate mental representation of text. Let us shortly illustrate this by means of an example.

(2)   [Context: Hans schaute aus dem Fenster. (*John looked out of the window.*)]
a.   Das Wetter war schön. (*The weather was nice.*)
b.   "Das Wetter ist schön.", dachte er. ("*The weather is nice." he thought.*)
c.   Das Wetter war doch schließlich schön. (approximately: *The weather was nice, after all.*)
d.   War das Wetter nicht schön? (*Wasn't the weather nice?*)

All four versions, (2a)–(d), share the same propositional content: that the weather is nice. But they differ considerably in two respects: first, who the holder of the propositional attitude is; and, second, what the relation between the holder and the attitude is. In (2a), the proposition could be entertained by the narrator, or by John; (2a) is, in this sense, ambiguous. This ambiguity is resolved in (2b) by the explicit factive component which attributes the proposition to John, and states that the relation between the holder (John) and the attitude (that the weather is nice) is a factive one, i.e. one of belief: John stands in the BELIEF-relation to a token of the proposition. In (2c), the same BELIEF-relation seems to hold between John and the proposition, but this comes about *without* the relation being explicitly expressed by a factive verb. Note that the proposition in (2c) can, under normal circumstances, not be attributed to the narrator; this seems to be an effect of both the syntactic features of FID (shifted tense and pronouns) plus the discourse particles added to the sentence. Finally, (2d), in which FID is marked by the interrogative mode, expresses a different relation between John and the proposition: while it certainly expresses that it is John and not the author who is entertaining the thought that the weather is nice, the relation is not necessarily one of belief, but could also be one of wondering: it does not seem to be outright factive. In any case, FID, then, is a means by which narrators can take on the perspective of a protagonist and relate details of the protagonist's mental life. It is this feature of *perspectivization* that may be taken to account for the pervasiveness of FID in modern literature: it allows to report the inner life of a protagonist without having the whole text to be narrated from that protagonist's first person perspective.

As noted above, the cues that mark a sentence, or discourse segment, as FID, are rather innocent: turning an assertion into a question, or adding a speaker-related particle (plus the obligatory tense shift) may suffice to cue the reader to switch to a FID interpretation. We assume that this interpretation involves a deviation from the canonic interpretation that a stretch of reported speech as in (2b) may get: the reader is invited to take on the perspective of the protagonist (to "get into his or her shoes") and to entertain the proposition from the protagonist's perspective. This means that all three components of the propositional attitude – the holder, the proposition, and the relation between them – are affected by marking a stretch of text as FID. In particular, the role the attitude holder plays is a special one: on the

one hand, he or she is mentally represented as a discourse referent, just as any other entity mentioned in the text. On the other hand, this discourse referent receives a special status by the fact that a certain state of affairs is related from his or her privileged perspective, and by the fact that the reader is invited to take over this perspective. Thinking of a discourse entity as an entry in memory, or, more technically, as a node in a network, it is not unreasonable to assume that the special status of that discourse referent is reflected in a higher activation state of the respective node in the network that stores and manages discourse referents and the relations between them. It is this special status that leads us to formulate our first hypothesis:

H1 The discourse referent representing the protagonist receives a higher activation in a discourse segment marked as FID, compared to one not marked as FID.

Given this hypothesis, the next step is to apply the methods of experimental psycholinguistics to test it. One obvious candidate for testing the activation of a discourse referent is anaphor resolution: if a discourse referent is highly activated, a pronominal expression referring back to it should be easier to resolve to that referent than when it is less activated (e.g. Sanford, Moar & Garrod 1988; Gundel 1993; Ariel 2001).

A few studies report that choice and interpretation of referential expressions are affected by point of view. Hewitt (1995) found that in subjective parts of stories, pronominal and zero anaphoric reference to the protagonist from whose perspective the passage is narrated is increased. Because the degree of activation of a referent at a certain point in a discourse is reflected in the form of the anaphoric expression used by the speaker to refer to that referent (Ariel 2001) and personal pronouns are exemplars of markers for highly activated referents, this finding is suggestive of an effect of narrative perspective on the activation status of the protagonist.

More recently, Kaiser and Cohen (2012) have shown that participants are sensitive to a rather subtle FID manipulation (similar to ours) when asked to indicate their anaphor resolution preferences on a 5-point scale. Sentences in the FID condition led to higher proportions of object choices than the "plain" condition. Interestingly, the effect of FID on these preferences was correlated significantly with the spatial perspective-taking abilities of the participants measured by the Perspective Taking/ Spatial Orientation Test (Hegarty & Waller 2004). Given these off-line findings, we were interested whether reading times – being a measure of on-line processing – would reflect the effect that FID exerts on the processing of pronouns.

In sum, we propose here that the protagonist whose point of view is presented receives a prominently high activation status, as pronouns are more often used to refer to that protagonist and the resolution preferences of pronouns are affected by FID. The potential effect of point of view, and more specifically FID, on ease of

anaphor resolution during on-line reading has, to our knowledge, so far not been tested in controlled experiments.

Returning to our reasoning about the role FID plays in the representation of propositional attitudes, recall that the protagonist, i.e., the attitude holder, is only one part of the characteristic interpretation that FID induces. The second one is the proposition itself. Again, it is not unreasonable to assume that, because the reader takes on the perspective of, or even identifies with, the protagonist, the proposition that the protagonist entertains is represented in the discourse model of the reader in a different way than a proposition which is merely related by the (omniscient) narrator. We eschew the question whether "represented in a different way" should be translated into a difference in activation status. But we think that it is safe to assume that the respective proposition will be encoded more deeply in the memory of the reader, and we are confident that the above reasoning warrants the formulation of the following hypothesis:

H2 A proposition marked by FID as being entertained from a protagonist's perspective should be easier to recall than a proposition not marked by FID.

One might even go so far and hypothesize that the linguistic form in which the proposition is expressed is easier to recall if the sentence is marked as FID. That is, an even stronger hypothesis than H2 might be put forward, which would claim that the recall for *verbatim* information is better with FID than without it. This hypothesis bears on discussions in the literature on the formal semantics of FID. There, one problem consists in the question whether the representation of FID shares properties with *direct discourse* (see e.g. the (mixed) quotation approaches put forward by Schlenker 2004 and Maier 2014), or whether FID is better analysed as a form of *indirect discourse*, as argued e.g. by Sharvit (2008). Not least because there is emerging psycholinguistic evidence that the mental representations of direct vs. indirect discourse can be differentiated by on-line measures (see Yao & Scheepers 2011), we hope that our data will shed light on the issues connected to H2.

We set out to test these hypotheses, i.e. that Free Indirect Discourse can be induced by certain linguistic markers, and that it affects discourse processing, and, more specifically, anaphor resolution, and recall. We tested them by measuring word-by-word self-paced reading times for discourse segments containing a long-distance antecedent-anaphor relation stretching across three sentences in the discourse (see Example (3) below), with the antecedent denoting the protagonist. We manipulated the presence and the type of FID indicator (i.e., indicator (particle or question) present (+ FID condition) vs. absent (− FID condition)) under the assumption that the presence of the indicator should result in a special status of the mental representation of the protagonist (H1), and, consequently, resolving the

anaphor to the antecedent should be easier if the indicator is present then when it is absent. Our main dependent (on-line) variable was word-by-word self-paced reading times on the sentence containing the anaphor; in addition, participants were presented with comprehension questions, which provided us with two further dependent (off-line) variables: percentage of correct answers to the comprehension questions, and the latency of the answer. For all three dependent variables, we hypothesized an effect of the factor FID (– FID vs. + FID), that is, word-by-word reading times on the anaphoric expression and the words following it should be lower, percentage of correct answers should be higher, and response latencies should be lower when the indicator is present than when it is absent.

In Experiment 2, we added a memory test after the on-line reading task. Participants had to judge whether a certain sentence had been part of the reading experiment they had just performed, and we recorded their reading times for the sentences, the latencies of their responses, and the correctness of their responses, although we only performed statistical tests on the percentage of correct responses. For these, we hypothesized that the proposition should be easier to recall if the sentence expressing it was marked as FID in the reading experiment, and thus was more deeply encoded (H2). This should result in a higher score of correct responses (hits, correct rejections) for the + FID condition as compared to the – FID condition. To our knowledge, the only experimental study testing for potential memory effects of FID was László (1986). We will report on his findings in the section where we introduce our own memory test.

3. Experimental evidence

3.1 Materials

We constructed 20 short narratives which prominently introduce one protagonist (alongside others) who is witness to a more or less prototypical event (a party, a stroll on a nice autumn day, a visit to concert etc.). Each narrative consisted of 8 sentences. Sentences 1–3 served to introduce the protagonist and the scenario; sentence 4 was the critical sentence which did or did not feature the FID cue. Sentences 5–7 contained descriptions of an event or situation which could either be construed as being narrated from the perspective of a neutral (authoritative) narrator, or from the perspective of the protagonist. Sentence 8 featured the protagonist in the form of an anaphoric expression ('er'/'sie' (*he, she*) in the role of grammatical subject, and contained two coordinated verb phrases. The coordination was included to provide us with an additional position to check for the effect

of FID, namely the position where the discourse referent of the protagonist (the elided subject, identical to that of the first conjunct) would have to be reactivated. An example of an experimental item with an approximate English translation is given in (3); the antecedent (= protagonist) and all anaphoric expression coreferent with it are <u>underlined</u>, the critical anaphoric expression is printed in **bold face**.

(3)   s1:   <u>Peter</u> schob sich an den Leuten im Flur vorbei in Richtung Küche.
      s2:   <u>Seine</u> Freunde waren schon um elf gegangen,
      s3:   aber <u>er</u> hatte beschlossen noch zu bleiben.
      s4:   Die Party war [∅/doch schließlich] in vollem Gange[./!]
      s5:   Die kleine Wohnung platzte fast aus allen Nähten.
      s6:   Auf dem Balkon drängelten sich die Raucher.
      s7:   Die Nachbarn hatten sich schon zweimal wegen des Lärms beschwert.
      s8:   <u>**Er**</u> ging zum Kühlschrank und griff sich noch ein Bier.

   (*Peter elbowed his way to the kitchen through the other people on the corridor. His friends had gone at 11 o'clock already, but he had decided to stay. [∅/After all], the party was in full swing[./!] The tiny flat was about to burst. The smokers were huddled on the balcony. The neighbors had complained about the noise twice already. He went to the fridge and grabbed another beer.*)

It was taken care that the overall structure of these narratives was uniform, since our main dependent variable, the reading time for words of the eighth sentence, which contains the anaphor, is known to be sensitive to the structural properties of the discourse sequence intervening between antecedent and anaphor (e.g. Holler & Irmen 2007). Thus, we made sure that the rhetorical skeleton of our narratives was identical across the experimental items. This skeleton is spelled out in (4) below in terms of the rhetorical relations (represented by the labeled arcs; "Cont." is short for "Continuation") holding between the propositions (represented by π_1 to π_8; e.g. Asher & Lascarides 2003)

(4) π_1 *Narration* π_2 *Contrast* π_3 *Narration* π_8

Explanation π_4

Elaboration

π_5 *Cont.* π_6 *Cont.* π_7
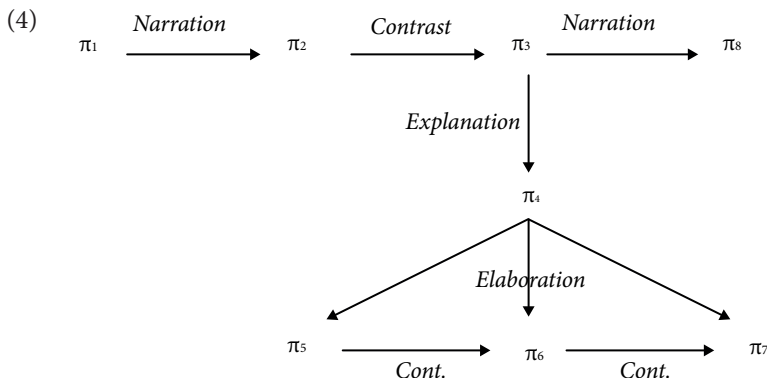
The proposition π_1, expressed by sentence s1, sets the stage for the events to be described in the course of the narrative. Sentence s2 expresses a proposition π_2 which stands in a local contrast to the next proposition, π_3 (in the example: there is an adversative relation between Peter's friends having left already and him having decided to stay nonetheless). The critical sentence s4, which either featured or did not feature the FID indicator, gives a reason for the proposition π_3 – the latter thus stands in an *Explanation* relation to the former. This reason is then elaborated on by the propositions expressed by sentences s5 to s7, which are related to each other by the default relation of *Continuation*. The final sentence s8 takes up the main line of events interrupted by s4 and continues the narration. The complete list of stimuli is available from the first author upon request.

3.2 Pilot 1: Off-line questionnaire study

Our first attempt to pin down the effect of FID cues was a 2 alternative forced-choice experiment, testing four of our overall 20 experimental items. 223 participants (undergraduates from the German Department and the Department of Psychology at Göttingen University) were presented with one item at a time in a small booklet and were asked to perform a forced choice decision task on the question whether the text described the events (a) from the perspective of a neutral narrator, or (b) from the perspective of the protagonist. Each participant got to see two of the four items in the + FID, and the other two in the – FID condition. For the + FID condition, we used the three types of variants: (i) particle 'doch' (literally: *though*), no exclamation mark; (ii) particle 'doch' plus exclamation mark; (iii) particle 'doch schließlich' (*after all*) plus exclamation mark. We created four different lists by assigning + FID condition to half of the four items and – FID to the other half; each type of FID variant was held constant within one list, i.e. one participant saw one type of FID variant only. Order of items was randomized.

We predicted that the proportion of decisions should be dependent on the FID factor: in the – FID condition, participants should judge the text less often as being told from the perspective of the protagonist as compared to the three + FID conditions. In addition, we wanted to check for an additional effect of the cue strength. Here, the prediction was that decisions for perspectivized narration should increase with markedness, and thus with cue strength (from (i), 'doch', to (iii) 'doch schließlich!').

37 participants had to be excluded from the experiment for technical reasons (2 were tested twice, nine were handed booklets in which conditions were mistakenly not varied across items, and 26 turned out to be non-native speakers of German). The data from the remaining 186 participants (130 female, 56 male;

mean age 21.4 years, $SD = 4.97$) were entered into the analysis. The binary variable (relative frequency of decisions for protagonist perspective) showed an effect of the FID condition (see Table 1 for the descriptive statistics). The data were submitted to a generalized linear mixed model with a logit link function (Barr, 2008) with participants and narratives as random factors, and FID condition as fixed factor. Following the recommendations of Barr et al. (2013), and thereby employing the most conservative test currently available, we included random slopes for items in our model. The factor FID had a significant overall effect on participants' decisions ($|z| = 5.88$, $p < .001$), but planned comparisons revealed that the three + FID variants (i)–(iii) above did not differ significantly from each other ($ps > .10$).

Table 1. Mean relative frequencies of decision for protagonist perspective depending on FID condition in Pilot study 1

	− FID	+ FID		
	∅	(i) doch	(ii) doch!	(iii) doch schließlich!
Proportion	.30	.62	.61	.70

The results of the first pilot study encouraged us to use the particle 'doch schließlich' for the reading time experiment as an indicator for FID, since it descriptively had the highest score of decisions for the protagonist perspective, and thus appeared to be best suited to induce a FID reading for the second part of the narrative.

3.3   Pilot 2: Expert rating

The overall set of 20 items was tested once more with 25 participants, who, in a sense, were experts on the interpretation of literary texts, since they were undergraduate and graduate students of literary studies. Participants as these may perhaps be expected to be more sensitive to the FID manipulation given their vast experience with the stylistic devices of literary texts, and thus to be especially qualified to detect differences between the experimental conditions. Be that as it may, four of them had to be discarded from the analysis because they were non-native speakers of German. Of the remaining 21 participants, 14 were female and 7 male; the mean age was 26, $SD = 1.45$. The overall quality of each of the 20 texts had to be assessed on three 7-point scales: (i) comprehensibility, (ii) plausibility, and (iii) linguistic form. In addition, participants were encouraged to mark passages in the text that they found particularly incomprehensible, implausible, or poor in linguistic form. For each participant, half of the items were presented in the condition + FID (with the FID indicator), and half of the items in the − FID condition. The factor FID was counterbalanced across two lists; items 1–12 featured the particle 'doch schließlich'

and the exclamation mark in the + FID condition, items 13–20 featured a rhetorical question in this condition. Since rhetorical questions are regarded as indicators of FID (see Eckardt 2014), we decided to add this indicator to the design and test for the relative strength of the two types of FID indicator (between items). The two lists were randomized. Two further lists were created by inverting the order of the first two lists (testing for order effects), yielding an overall of four lists.

We predicted there to be no differences between the texts with respect to their comprehensibility, plausibility, and linguistic form. The results of the comprehensibility, plausibility, and form ratings (1 = poor, 7 = good) are given in Table 2, separately for the two FID conditions. Items in the + FID condition were rated somewhat better on all three scales. We computed by-subject and by-item ANOVAs on these data, as well as linear mixed effects models (LMMs) with maximal random effect structure (Barr et al. 2013). Although the comprehensibility difference was significant by subjects in the ANOVA ($F_1(1,20) = 4.60$, $p < .05$), it failed to reach significance in the by-items ANOVA ($F_2(1,19) = 3.76$, $p = .07$; similarly for the LMM ($LR_{\chi 1}^2 = 3.49$, $p = .06$). None of the remaining differences reached significance in ANOVAs or LMMs, all ps > .10. The difference between the + FID and the − FID condition means for each item never exceeded 1.5 points on the scale.

Table 2. Mean ratings (7-point, 1 = poor, 7 = good; standard deviations in brackets) depending on FID condition in Pilot study 2

scale	− FID		+ FID	
comprehensibility	5.96	(1.07)	6.13	(.89)
plausibility	5.36	(1.35)	5.57	(1.38)
linguistic form	4.72	(1.31)	4.80	(1.25)

Items which received overall low ratings were checked by the authors, and possibly incomprehensible, implausible or otherwise awkward pieces of text were corrected. To enhance the overall comprehensibility and plausibility, headings were added to the texts. In the example above, the heading was "WG-Party", which could be translated as "Party at a flat-sharing community".

3.4   Experiment 1: Self-paced reading

The overall set of 20 items was intermixed with 16 items from a different experiment concerned with the interaction of discourse structure and information structure (4 conditions), yielding a set of 36 texts overall. One training item was construed along the lines of the text skeleton laid out above. Four lists containing the 36 experimental items were created; the FID conditions were counterbalanced across

the four lists, as were the condition of the filler experiment. The four lists were assigned to participants according to a latin square design. Each experimental list was randomized for each participant.

3.4.1   *Experiment 1: Procedure*

Participants ($N = 27$, 5 male, 22 female, mean age 23.9, $SD = 4.34$) were tested individually at a PC workstation. The LINGER experimental software (version 2.94, written by Doug Rohde, available at http://tedlab.mit.edu/~dr/Linger/) was used to display the stimuli. Participants were seated comfortably in front of the keyboard and monitor. The instruction was displayed on the first screen, which told participants that they would be reading short texts in the following fashion: the first six sentences of each text (including the headings) were presented one sentence at a time. All eight sentences of the text were displayed as underscores, with blanks indicating word boundaries. Participants were told to read the sentences carefully, since they would be asked comprehension questions after each text. As soon as they had understood the sentence, they were instructed to press the space bar on the keyboard to display the next sentence on the monitor, while the previous sentence disappeared (non-cumulative sentence-by-sentence self-paced reading). From sentence 7 on, presentation was word-by-word (non-cumulative word-by-word self-paced reading). Participants were instructed that they would have to press the space bar to display the words of the sentences one by one, and that the previous word would disappear, and that they, the unnatural mode of presentation notwithstanding, should try to approximate their natural reading speed. After the word-by-word reading of sentence 8 was completed, pressing the space bar displayed the comprehension question (yes/no-questions). Participants were asked to indicate their response by pressing the keys for the numbers "1" on the number block of the keyboard for "yes" and "3" for "no". Questions asked for information from the first half of the item (in our example, this could for example be the question whether Peter had decided to leave the party), or for information encoded in the last sentence (in our example, the question was: "Did Peter grab a coke?"). The pairing of question and answer was fixed for each item. After the response to the comprehension question was given, participants were presented with a 'PAUSE' screen and were asked to press the space bar to begin the presentation of the next item.

Participants were told that there would be a training phase after which any further questions or problems would be dealt with. In the training phase, there were two trials habituating the participants to the word-by-word reading procedure, the second of which was followed by a comprehension question; the third trial was a text with eight sentences, with sentence-by-sentence presentation for sentences 1–6, and word-by-word presentation for sentences 7 and 8 (as in the experimental trials). After the training phase, any questions the participant would still have were

answered by the experimenter. When the participant had no further questions, s/he would start the experiment. Overall, the experiment took approximately 35 minutes.

3.4.2   *Experiment 1: Design and predictions*

The design consisted of the 2-level factor FID (– FID: no FID cue present in text; + FID: FID cue present), realized within participants and items. Given the hypothesis in Section 2, our main prediction was that the reading times for the anaphoric element in the target sentence (sentence 8), and on the material immediately following it, should be lower in the + FID condition than in the – FID condition, since the anaphor should be resolved faster in this condition due to the hypothesized higher activation level of the discourse referent denoting the protagonist. In addition, the design contained two within-subjects, but between-item factors: the type of FID cue (FIDTYPE, particle vs. question), and question type (QTYPE, comprehension question pertained to an early part of the passage, i.e. before the FID manipulation had occurred in the text, or to the last sentence, which was, so to speak, in the scope of the FID cue). The QTYPE factor was included for exploratory purposes. Regarding this factor, we hypothesized that questions pertaining to the last sentence should show higher accuracy rates, and shorter response latencies, than questions pertaining to an earlier part of the text. In addition, we predicted that both accuracies and latencies would show an effect of the presence of the FID cue: when it is present, accuracies were predicted to be higher, and response times shorter, since we hypothesized that the passage was encoded more deeply due to perspectivization.

In addition to testing our hypotheses about the processing of FID, we wanted to compare three different ways of analyzing data from repeated measure designs like the one employed in our main experiments: (i) mixed model ANOVAs with participants and items as random factors; (ii) linear mixed models with random intercepts for participants and items; and (iii) linear mixed models with both random intercepts and random slopes for participants and items. Given the recent claims about differences in conservativity between these methods (cf. Barr et al. 2013), we were interested which of these methods would be able to detect a potential effect of FID in our data, and how the different methods would perform in comparison to each other. Here, especially the difference between (ii) and (iii) would be potentially interesting given that only (iii) could account for potential interactions between particular items, participants, and the experimental manipulation, which is a problem that especially data coming from reading larger stretches of text might be beset with. This might be traced back to the fact that items consisting of multiple propositions which constitute a scenario are probably more prone to show interactions with the world knowledge the individual participants bring to the experiment, thereby producing a larger amount of variance in the data. The reason for choosing

these three types of statistical models was made simply on the grounds that these are currently the ones that seem to be most frequently applied by researchers in psycholinguistics.

3.4.3 Experiment 1: Results

Data treatment

Experiment 1 provided us with three dependent variables: reading times per word for the critical sentence; proportion of correct answers to the comprehension questions; and response latencies for the answer to the comprehension question. Reading times for trials in which the participants gave an incorrect answer to the comprehension questions were discarded. This led to an exclusion of 331 data points (6.3%). Word-wise reading times were log-transformed and observations beyond 2.5 standard deviations from the participant mean were excluded from further analyses; this affected further 100 data points (1.9%). Furthermore, response latencies for trials with incorrect answers (31 data points, 5.7%) and response latencies above 5000 ms were discarded (33 data points, 6.1%).

On the reading times and the response latencies, we performed three types of analyses: classical mixed model ANOVAs with participants and items as random effects (Analysis 1); linear mixed models with participants and items as random effects, but with random intercepts only (Analysis 2); and maximal linear mixed models with participants and items as random effects and additional random slopes (Analysis 3; Barr et al. 2013).

The proportions of correct answers were submitted to linear mixed models with a logit link function (Jaeger 2008).

Proportion of correct answers

The mean percentages of correct answers to the comprehension question are given in Table 3.

Table 3. Mean proportion of correct answers (and SDs) to the comprehension question in Experiment 1, dependent on FID, QTYPE, and FIDTYPE.

FIDTYPE	QTYPE	− FID	+ FID
particle	early	.91 (.10)	.91 (.08)
	late	.96 (.09)	.98 (.09)
question	early	.91 (.11)	.91 (.07)
	late	.98 (.07)	.98 (.08)

As a quick glance at Table 3 reveals, there were no effects of FIDTYPE or QTYPE on the proportion of correct answers, apart from the trivial effect of QTYPE (early vs. late).

Reading times per word in target sentence

The mean reading times per word in the target sentence are illustrated in Figure 1. As noted above, we submitted the reading times[1] to three different analyses. However, none of these revealed a significant effect of our experimental manipulation: Analysis 1: all Fs < 1, both in the by-participants and the by-items analysis; Analysis 2: all $|t|$s < 1; dito for Analysis 3. This held both for the analysis of raw reading times (which we plot in Figure 1 for perspicuity), as well as for log-transformed reading times.
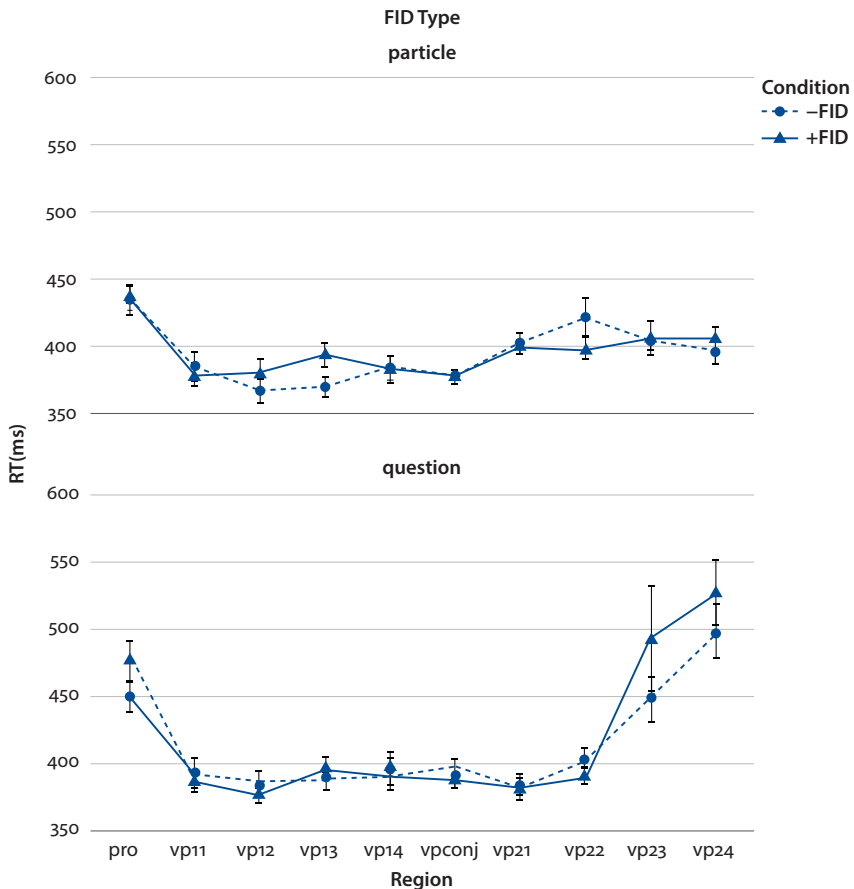
Figure 1. Raw reading times per word for the critical sentence dependent on FID and FIDTYPE for Experiment 1; error bars depict 1 standard error of the mean. Plotted on the x-axis are the regions of the critical sentence ("pro": pronoun; "vp11–vp14": constituents of first verb phrase; "vp1con": conjunction; "vp21–vp24": constituents of second verb phrase.

1.   For each sentence the reading times for the pronoun, the first two words following the pronoun, the conjunction and the first two words following the conjunction were analysed.

Response latencies

The mean response latencies for the answers to the comprehension questions are illustrated in Figure 2 below.
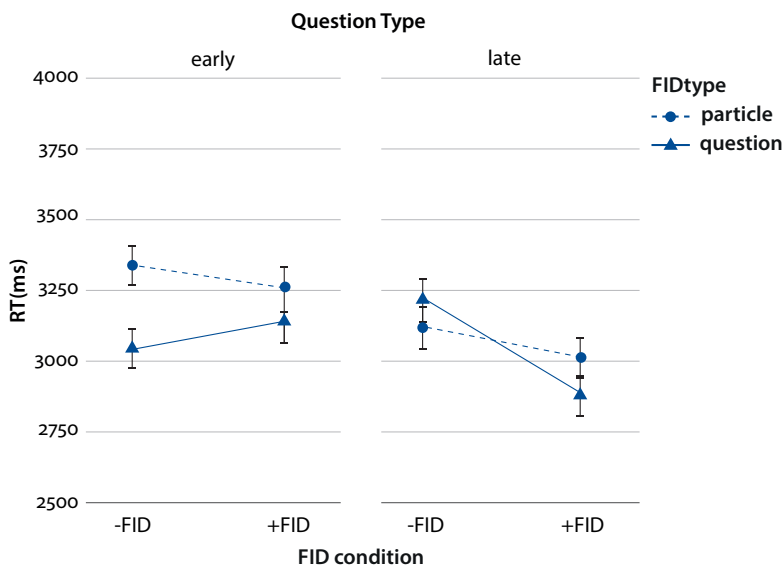
Figure 2. Response latencies for answers to comprehension questions dependent on FID (see x-axis labels), FIDTYPE (circles = particle; triangles = question) and QTYPE (left panel: question to early part of text; right panel: question to last sentence) for Experiment 1; error bars depict 1 standard error of the mean.

Here, our different analyses provided us with different results. Analysis 1, the mixed model ANOVA, yielded a significant main effect of QTYPE in the subject, though not in the item analysis: questions pertaining to the last sentence were answered faster than questions pertaining to an earlier sentence, 3025 vs. 3160 ms, ($F_1(1,26) = 6.46$, $p < .05$; $F_2(1,16) = 1.39$, $p = .26$). Moreover, Analysis 1 revealed an interaction – which, again, was significant by subjects only – of FID and QTYPE: whereas reaction times were slightly faster in the $+$ FID versus the $-$ FID condition for questions to early sentences (3141 vs. 3179 ms), for questions to late sentences a substantial effect of FID emerged (3115 ($-$ FID) vs. 2937 ms ($+$ FID); $F_1(1,26) = 5.23$, $p < .05$; $F_2(1,16) = 1.27$, $p = .28$). Also, the three-way interaction between FID, FIDTYPE and QTYPE turned out significant in the subjects, but not in the items analysis ($F_1(1,26) = 4.86$, $p < .05$; $F_2(1,16) = 1.72$, $p = .21$).

Going on to Analysis 2, the linear mixed model exhibited no significant main effects or interactions, all $|t|s < 2$. The same holds for Analysis 3, the linear mixed model with maximal random effects structure.

3.4.4   *Experiment 1: Discussion*

The interaction of FID and QTYPE which the by-participants ANOVA reported to be significant seemed to us to make a lot of sense: since the FID manipulation occurred *after* the sentence queried in the early QTYPE condition, it should not show any effect, while questions asking for information about the last sentence could, and, descriptively, indeed did show an effect of the FID manipulation. However, none of the effects which the ANOVA procedure had flagged as significant turned out to withstand when we performed tests that take into account the combined variance exerted by participants and items, as well as their possible interactions with the fixed effects.

Given these mixed results, we were wondering where the differences for the different procedures may come from. Given the lack of a significant result in the by-items-ANOVA, an obvious place to look was the variance introduced by our experimental items. Inspection of the 95% confidence intervals by item showed that the main effect driving the interaction (the effect of FID in the late QTYPE condition) was significant for one item only (item 17). To make sure that this item was the culprit, we calculated Cook's distance using the influence.ME package for the R statistics software (Nieuwenhuis, Pelzer & te Groetenhuis 2012). This package provides a procedure which iteratively removes observation units from the data set and checks whether this omission of units has an effect on a given model. It outputs an estimate of the influence of (groups of) data points on the overall pattern, Cook's distance, abbreviated as D_i, which can be used to pinpoint particularly influential data points and has been argued to have a cut-off point at $D_i > 4/n$, where *n* is the number of observation units. The only item whose D_i was beyond that cut-off point was, again, item 17, whose D_i was larger than 4/20 – in fact, it was even larger than 0.4.

From the overall results of Experiment 1, a number of interim conclusions can be drawn: firstly, there seems to be no on-line effect of FID on anaphor resolution processes in self-paced reading. Moreover, the effects of the FID manipulation on off-line measures like correctness and latencies of answer to comprehension questions seem to be rather subtle, and apparently do not withstand closer statistical scrutiny. Thirdly, the results from the by-item analysis, together with the finding that one of our items exceeded the criterial value of Cook's distance and thus seemed to be carrying the effect almost single-handedly, suggested to us rather unequivocally that our materials were suboptimal and that we should have a second go at our items. Experiment 2 was an attempt to remedy the defects detected in Experiment 1.

3.5 Experiment 2: Self-paced reading and memory test

We conducted a follow-up experiment to rule out the possibility that the absence of an effect of Free Indirect Discourse on reading times in our first study was due to the prominent position of the sentence that contained the long-distance anaphor to the protagonist within the narratives (it was always the last sentence). Therefore an additional coda sentence was appended to the end of each narrative.

In the first experiment the inferential statistics revealed a mixed picture for the off-line-effects. Most of the comprehension questions in the previous experiment had a low level of difficulty. To test if a ceiling effect may have obscured possible effects of FID on answer variables to comprehensions question in the first study, we revised the comprehension questions and made a substantial part of them less easy.

Additionally, a second off-line measure, a recognition test for the critical sentences from the narratives, i.e. the sentences of each narrative that were manipulated for presence or absence of FID cue, was supplemented to the experimental procedure. László (1986) reports that participants show better sentence recognition performance with phrases written from, as he calls it, an internal point of view (i.e., from the protagonist's perspective) than with sentences written from an external point of view. Participants in László's study seemed to have remembered the point of view from which a phrase was written and could therefore reject a paraphrase with differing point of view more often for passages written from the perspective of a character. Kintsch (1998: 206) remarks that for memory and comprehension of literary texts, surface structure should play a bigger role than for nonliterary texts. In general, memory for exact wording of sentences decays rapidly (e.g. Bransford, Barclay & Franks 1972; Garnham 1981), but for sentences that express evaluative information of the speaker and are more representative of oral discourse, memory for verbatim information is enhanced (Long 1994). Because Free Indirect Discourse is a literary method for representing a character's speech and thought from an internal point of view, we hypothesized that markers of FID would evoke better memory for the exact wording of sentences (cf. our hypothesis H2 above).

3.5.1 *Materials*
Some modifications were carried out on the stimuli of our previous experiment. An additional sentence was appended to the end of each narrative as you can see in (5), which is an example item from Experiment 2.

(5) s1: Peter schob sich an den Leuten im Flur vorbei in Richtung Küche.
     s2: Seine Freunde waren schon um elf gegangen,
     s3: aber er hatte beschlossen noch zu bleiben.
     s4: Die Party war [∅/doch schließlich] in vollem Gange[./!]

s5:   Die kleine Wohnung platzte fast aus allen Nähten.

s6:   Auf dem Balkon drängelten sich die Raucher.

s7:   Die Nachbarn hatten sich schon zweimal wegen des Lärms beschwert.

s8:   **Er** drang zum Kühlschrank vor und griff sich noch ein Bier.

s9:   Jemand hatte zum dritten Mal Nirvana aufgelegt.

(*Peter elbowed his way to the kitchen through the other people on the corridor. His friends had gone at 11 o'clock already, but he had decided to stay. [∅/After all], the party was in full swing[./!] The tiny flat was about to burst. The smokers were huddled on the balcony. The neighbors had complained about the noise twice already. He forced his way to the fridge and grabbed another beer. Someone put on a record by Nirvana for the third time.)*

For nine narratives, target sentences (for which reading times effects of FID were predicted) were modified in order to achieve a more similar number of words for those sentences between the narratives. The changes to the target sentences concerned the addition of adjectives and adverbs and/or the modification of verbs set to describe an action the protagonist performed. Apart from one sentence, the meaning of the sentences was not altered substantially by these changes. Also, 15 of the total of 20 comprehension questions were made slightly more difficult. For five questions, the type of question was changed in comparison to the first experiment, resulting in a total of 11 questions of type 'late' und 9 questions of type 'early'.

During the recognition test, half of the critical 20 sentences of the narratives were shown in the same version (match) and half in the other version (mismatch) as the participants had seen them in the self-paced reading experiment. Thus, the correct answer to half of the experimental items in the recognition test would be 'yes' and to half of them of them 'no'. Filler items in the recognition test consisted of 36 items from the reading phase and 40 newly constructed sentences, which resulted in a total of 20 experimental and 66 filler items. The filler items from the reading phase consisted of 20 sentences other than critical sentences from the experimental narratives (e.g. *The smokers were huddled on the balcony*) which required a 'yes' response and 16 sentences from the filler items, with half of the sentences requiring a 'yes' and half a 'no' response. The remaining 40 "new" filler sentences were not part of any of the narratives nor paraphrases of sentences shown during the reading phase.

3.5.2   *Procedure*

The self-paced reading procedure was exactly the same as in the previous experiment. Immediately after completing the self-paced reading task, participants' recognition memory for the critical sentences from the narratives (FID cue present vs. absent) was tested. Before the recognition test started, participants received

written instructions about the procedure. Participants were told that their task was to decide as fast and accurately as possible whether the test-sentences had been part of the self-paced reading experiment they just completed. A training phase without feedback followed the instructions to enable the participants to familiarize with the procedure. Sentences were displayed as underscores, with blanks indicating word boundaries. The test-sentence appeared by means of a press of the spacebar. A second press of the space bar led the sentence to disappear and the question "Was this sentence presented in this way during the experiment?" was shown on the screen as well as the mapping of the response keys (3) and (1) to the respective answers 'yes' and 'no'. No feedback about the correctness of the answers was given.

3.5.3 Participants
27 participants were tested for the follow-up study. All were students of Göttingen University and attended the experiment for course credits or monetary compensation. Mean age was 23 years (SD 2.4); of the 27 participants only those with German as their native language (24 participants, 6 male) were included into the data analysis.

3.5.4 Experiment 2: Design and predictions
The design for the self-paced reading experiment was the same as in Experiment 1. Again, our prediction was that reading times for the pronoun in the target sentence (sentence 8), and on subsequent words should be lower in the + FID condition than in the − FID condition. For the factor question type (QTYPE, early vs. late) we predicted that questions pertaining to the last sentence should show higher accuracy rates, and shorter response latencies, than those querying about an earlier part of the text and that accuracies for late questions should be higher, and response times shorter when the FID cue was present than when it was absent. Concerning the recognition test a higher proportion of correct answers was predicted for sentences that had been shown in the + FID condition during the reading phase than for sentences that had been presented in the − FID condition during the reading phase.

3.5.5 Experiment 2: Results
Data treatment
For the analysis of the data four participants were excluded, because their proportion of incorrect answers to the comprehension questions exceeded 15 percent. Thus, data from 20 participants entered the statistical analysis of reading times and

response latencies. As was done for Experiment 1, reading times for trials in which the participants gave an incorrect answer to the comprehension questions were discarded. Word-wise reading times were log-transformed and observations beyond 2.5 standard deviations from the participant mean were excluded from further analyses. This resulted in the exclusion of 692 data points (16.8%) in Experiment 2.

Response latencies for incorrect answer (61 data points, 15.3%) and those above 8000 ms (13 data points, 3.3%) were excluded.

We performed the same three types of analyses on the data as in Experiment 1. For the analysis of the proportion of correct answers in the recognition test, one item had to be excluded due to incorrect labeling in the experimental program. Unfortunately, the labeling error resulted in an unequal distribution of items across the conditions 'match' (n = 12) and 'mismatch' (n = 7) in the memory task. Thus, data from 19 items entered the statistical analysis of the item recognition test. The proportions of correct answers from comprehension questions and recognition test were submitted to linear mixed models with a logit link function (Jaeger 2008).

Proportion of correct answers

No significant effect of FID condition was observed on proportions of correct answers to the comprehension questions in Experiment 2 ($|z|$-values for all but one comparison are < 1, for qtype $|z|$ = 1.27, p = .20; for descriptive statistics see Table 4).

Table 4. Mean proportion of correct answers (and SDs) to the comprehension question in Experiment 2, dependent on FID, QTYPE, and FIDTYPE.

FIDTYPE	QTYPE	− FID	+ FID
particle	early	.90 (.16)	.93 (.14)
	late	.78 (.20)	.77 (.19)
question	early	.89 (.41)	.93 (.24)
	late	.79 (.49)	.79 (.50)

Reading times per word in target sentence

As in the previous study, presence or absence of FID had no effect on pronoun reading times or on reading times for subsequent words in the target sentence. This was true for all three analyses, that is, all *F*- and all absolute *t*-values < 1; see Figure 3 for illustration.
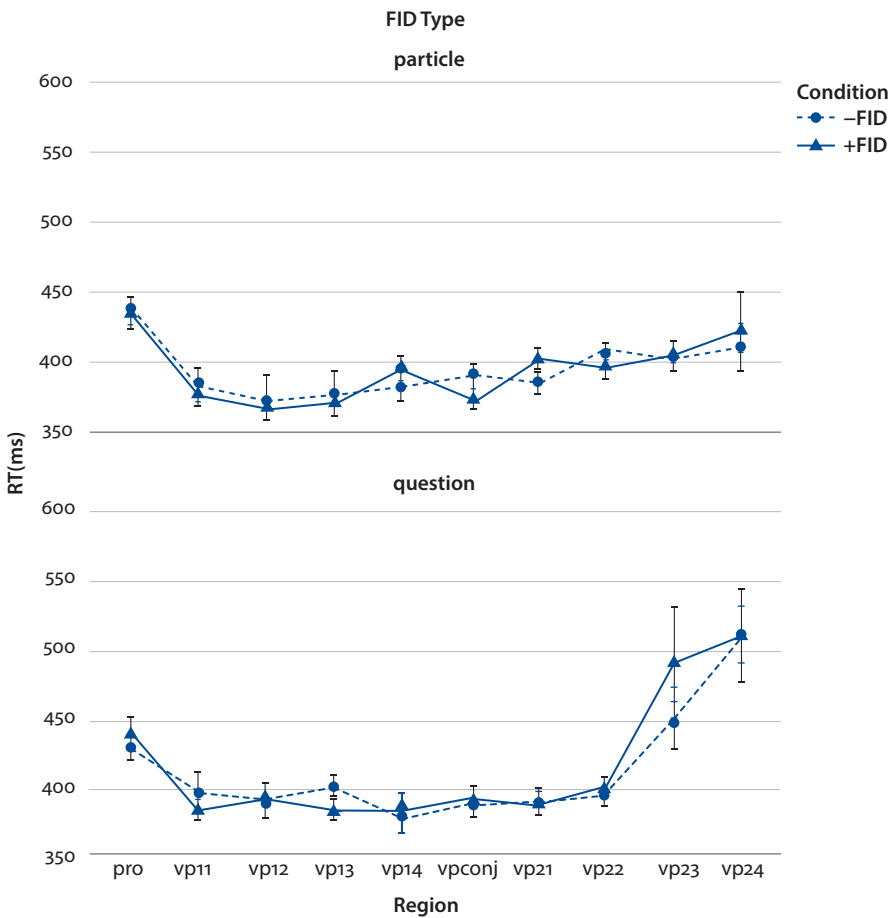
Figure 3. Raw reading times per word for the critical sentence dependent on FID (circles vs. triangles) and FIDTYPE (upper vs. lower graph) for Experiment 2; error bars depict 1 standard error of the mean.

Response latencies

In contrast to Experiment 1, all three analyses that we performed on the response latencies for the answers to comprehension questions converged on the same result: we found no effect of FID condition (descriptively, the effect even went into the wrong direction (see Figure 4). Main effects and interactions of the factors FID condition and question type were not significant.
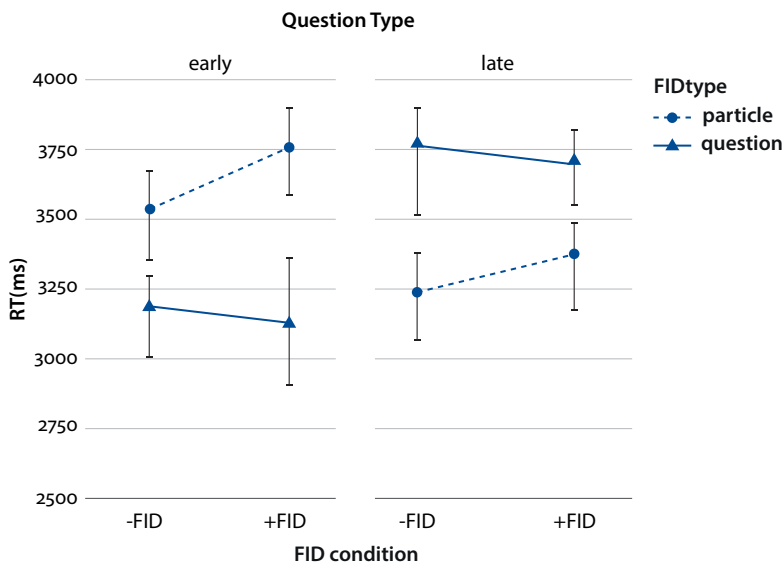
Question Type

Figure 4. Response latencies for answers to comprehension questions dependent on FID (see x-axis labels), FIDTYPE (circles = particle; triangles = question) and QTYPE (left panel: question to early part of text; right panel: question to last sentence) for Experiment 2; error bars depict 1 standard error of the mean.

Memory test

In the recognition task, which followed the self-paced reading part of the experiment, participants had to indicate whether a sentence shown to them on the computer screen had been presented in the same form in the preceding self-paced reading task or not. Accurate performance in this task requires the participants to both correctly accept matches and correctly reject mismatches.

Presence of FID marker within the critical sentence from the reading phase had a significant influence on overall recognition performance ($|z| = 2.29$ 2.5; $p < .05$), but this effect was not in the predicted direction (see Table 5 for descriptive statistics). Overall, participants showed better recognition performance for those critical sentences from the reading task that did *not* contain a FID marker (− FID: 72% correct answers) than for those with FID marker (+ FID: 58% correct answers).

Further exploratory analyses revealed that the difference could be attributed to a high rate of false alarms (incorrectly accepted mismatches) for those cases where participants were probed with the non-FID version of a sentence in the recognition task when they actually had seen the + FID-version of the respective sentence during the reading task (81% false alarms). The false alarm rate for the opposite

Table 5. Mean proportion of correct answers (and SDs) to the recognition test dependent on FID and FIDTYPE for Experiment 2; numbers in bold indicate hit rates and numbers in italics correct rejections.

	FID type			
	particle		question	
Reading Phase (target)	**Recognition Test (probe)**			
FID condition	**– FID**	**+ FID**	**– FID**	**+ FID**
– FID	**.82 (.22)**	*.42 (.46)*	**.95 (.15)**	*.52 (.38)*
+ FID	*.14 (.28)*	**.80 (.22)**	*.22 (.30)*	**.82 (.29)**

constellation was considerably lower (– FID version in reading phase and + FID version in recognition task: 52% false alarms). Hit rates (correct acceptances) differed only marginally between the two FID conditions (+ FID: 81%; – FID: 87%).

A significant interaction between the factors FID CONDITION IN THE READING TASK and FID CONDITION IN THE RECOGNITION TASK for the proportion of correct answers emerged ($|z| = 4.29; p < .001$). Two-sided *t*-Tests comparing the proportion of false alarms to chance level (50% false alarms) revealed a significant difference for sentences presented in the + FID condition during the reading phase ($|t|(68) = 6.57$, $p < .001$) whereas the false alarm rate for the respective – FID condition did not differ significantly from chance ($|t|(70) < 1$).

In summary, there was a general tendency to produce false alarms in the mismatch conditions, but this tendency was more pronounced when the sentence to be remembered had contained a FID cue. Thus the overall worse recognition performance for critical sentences with FID markers was mainly due to the high rate of false alarms that participants produced when they were probed for a sentence containing a FID marker from the reading task with the respective non-FID version of the sentence during the recognition task.

3.5.6 *Experiment 2: Discussion*

As in the previous study, Free Indirect Discourse did not influence reading times of the critical sentence containing the long-distance anaphoric relation to the protagonist. In the follow-up experiment, the sentence containing the long-distance anaphoric relation to the protagonist was not anymore in the final position within narratives, so the possibility that the missing influence of FID on reading times in the previous study was simply due to the fact that the critical sentence was in the final position within stimulus texts could be ruled out. Thus, the results from the follow-up experiment, like the first study, suggest that Free Indirect Discourse does not affect on-line anaphor resolution, at least for texts in which only one

main protagonist is introduced, and when measured by word-by-word self-paced reading. The data from the follow-up study also clearly showed that Free Indirect Discourse markers do not affect how often or how fast a correct answer is given to questions. This was true despite of the fact that parts of the comprehension questions were less easy than in our previous study. Thus, in the follow-up experiment we found no indication that Free Indirect Discourse leads to a different status of the protagonist in the mental model of the reader during text comprehension. Together with the results from our first experiment, this null result makes it rather hard to assess the validity of our hypothesis H1.

Our second hypothesis, H2, stating that Free Indirect Discourse markers and the sentences they are included in are processed differently than the same sentences without Free Indirect Discourse markers, was tested by the recognition test. The prediction that verbatim information is easier to recall when Free Indirect Discourse cues are present than when they are absent was not supported by the data. Overall recognition performance was even better for sentences without FID cues than for those with FID cues. Discrimination performance in general was poor for both conditions, but recognition errors occurred most frequently when the mismatch between test sentences and sentences from the reading phase was such that the sentence in the reading phase had contained a FID marker and the test sentence was shown in the alternate version without FID marker. This conflicts with the results of László (1986), who found better recognition memory for sentences written from an internal viewpoint than for those written from an external point of view. A possible explanation for this is that in László's study recognition probes from an external perspective were constructed to be inconsistent with an internal perspective, which seems to us to constitute a considerable difference to the probes we used in our study. Sentences without FID markers do not explicitly contradict an internal viewpoint, and therefore just remembering from which perspective a phrase was written may not have been as much an advantage in our experiment as in the experiment reported by László. Also, the FID markers we used may not belong to the class of expressions that evoke better verbatim recall. A possible post hoc explanation for the results of the recognition test could be that Free Indirect Discourse emphasizes the relevance of the content of a sentence to the goals of the protagonist and the whole mini-story. Therefore, whether sentences contained Free Indirect Discourse markers or not had an influence on how important the content of a particular sentence was perceived within an item. This is to say that the content of sentences with FID markers was perceived as more central to the situation described within the narrative than the content of the more neutral sentences without FID cues. Thus, it seems perfectly possible that participants perceived a higher familiarity for the propositional content of sentences that had contained a FID cue

and thus were more prone to accept sentences without FID cue as one they had seen before, even though it has not been presented in the same version during the reading phase. While we find this explanation for the pattern quite plausible, it is faced with the problem that the reasoning behind it makes a prediction not fulfilled by our data: it predicts a main effect of FID, which we did not find. Currently, we do not know how this dilemma should be solved. But we are confident to elucidate the relation between memory performance and FID by further experiments.

4.   General discussion

The results of the two experimental studies reported here give no indication for the validity of the hypothesis that narrative perspective mediated through Free Indirect Discourse influences saliency of protagonist referents, or of their mental attitudes, during on-line text comprehension. Reading times for anaphoric expressions refer-ring to the protagonist, which are indicators for referent accessibility in a number of contexts, were not influenced by the presence of Free Indirect Discourse. It could be the case that other factors than presence or absence of Free Indirect Discourse influenced the speed of anaphor resolution in the experiments, e.g. the fact that there was only one referent introduced as the main protagonist of the narratives or the discourse structure of the sentences following the critical sentence. This should be tested in future studies.

Taken together, the results from the two experiments reported here make the following conclusions very probable: firstly, our hypotheses about the on-line effect of FID on anaphor resolution cannot be upheld in the face of our current findings. Second, even the off-line effects of the FID manipulation that we chose (particles and polarity questions) seem to be rather subtle, which may partly be attributed to variance intruding via the experimental items. Third, given this susceptibility of the effects to spurious variance that the study of narratives has to face quite often, researchers should put their data to statistical tests which allow us to cope with these kinds of variances, while at the same time being as conservative as possible.

Our failure to find on-line effects may be attributed to a number of different reasons. Given the high amount of between-items variance that we reported for Experiment 1, and which may have carried over to the (albeit revised) items of Experiment 2, we suspect that the complexity of the narratives we presented to our participants may have induced variance which possibly has blurred the effects of our experimental manipulation. One aspect of the complexity is the relatively long distance between antecedent and anaphor. We had constructed the items in this way because we wanted to avoid any recency effects, and wanted the potential effect of FID to have some time to unfold. But note that the cues for FID were

relatively sparse: in the + FID condition, the critical sentence either contained two speaker-related particles, or was transformed into a polarity question. None of the sentence following the critical sentence contained any FID cues. But if one looks at passages construed as FID in literary texts, cues for FID are rather pervasive throughout that passage, i.e. they are frequently repeated and/or varied throughout that passage, potentially to remind the reader of the marked interpretation that the passage is to have. Our experimental items clearly lacked this kind of cue repetition, and that may have further affected the effect of FID. We are currently preparing an experiment in which we try to take these difficulties into account, e.g. by repeating FID cues more frequently, reducing the distance between antecedent and anaphor, and keeping the overall discourse structure of the narratives less complex.

The second major conclusion to be drawn from our findings is that the results from classical by-subjects ANOVA have to be taken with caution in designs with between-item factors, and (possibly) considerable between-item variance. This, of course, has been known since Clark (1973). Still, we think that this general problem of creating variance by varying the lexical material in items is particularly problematic for studies of narratives. For one thing, narratives per definition consist of multiple sentences, and are highly dependent on a number of possibly deleterious variables, like inter-individual differences in world knowledge, literacy, exposure to literary texts, to name only a few. It should be noted that the common practice of experimentally manipulating pieces of actual literary texts (as, for example, László did in his 1986 study, and many others after him: Long (1994), Bray (2007), Kotovych et al. (2011), to name a few) makes this point even more problematic. In addition, these inter-individual differences may interact with item-level properties, as well as with fixed effects. We think that the statistical procedures we have employed here, especially linear mixed effects models with maximal random effects structure, might be a remedy for that problem, although they surely are no compensation for careful item construction. Furthermore, we think that the differences between the results produced by the ANOVA vs. the linear mixed effects model nicely exemplify the critical point made by Simmons et al. (2011) whether an effect is significant or not is dependent, among other things, on the statistical procedure employed. It is a general rule that the most conservative tests should be employed. But given the current situation where multiple procedures for performing inferential statistic tests are available in parallel, it seems advisable to perform multiple analyses, and report their respective results. In addition, tools for plotting between-item and between-subject variance, as the influence.ME package mentioned in Section 3.4.4 above can be helpful in assessing the interplay of fixed and random effects beyond mere significance of results.

To sum up: we are still confident that on-line effects of FID can be found if all of the above problems are dealt with (or at least as many as possible), and we will

pursue the line of research started here in further experiments. But we think that the lesson to learn from the experiments reported here is that in the experimental investigation of narratives, both the experimental setup, as well as the statistical procedures employed, deserve a lot of care.

Acknowledgements

References

Ariel, M. 2001. Accessibility theory: An overview. In *Text Representation: Linguistic and Psycholinguistic Aspects* [Human Cognitive Processing 8], T. Sanders, J. Schilperoord, & W. Spooren (eds), 29–87. Amsterdam: John Benjamins.

Asher, N. & Lascarides, A. 2003. *Logics of Conversation*. Cambridge: CUP.

Banfield, A. 1973. Narrative style and the grammar of direct and indirect speech. *Foundations of Language* 10(1): 1–39.

Banfield, A. 1982. *Unspeakable Sentences*. Boston MA: Routledge and Kegan Paul.

Barr, D. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* 59(4): 457–474.

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.

Bortolussi, M. & Dixon, P. 2003. *Psychonarratology: Foundations for the Empirical Study of Literary Response*. Cambridge: CUP.

Bransford, J. D., Barclay, J. R. & Franks, J. J. 1972. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology* 3(2): 193–209.

Bray, J. 2007. The 'dual voice' of free indirect discourse: A reading experiment. *Language and Literature* 16(1): 37–52.

Clark, H. H. 1973. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior* 12(4): 335–359.

Eckardt, R. 2014. *The Semantics of Free Indirect Discourse. How Texts Allow to Mindread and Eavesdrop*. Leiden: Brill.

Fludernik, M. 1993. *The Fictions of Language and the Languages of Fiction*. Oxford: Taylor & Francis.

Garnham, A. 1981. Mental models as representations of text. *Memory & Cognition* 9(6): 560–565.

Genette, G. 1994. *Die Erzählung*. Stuttgart: UTB.

Gundel, J. K., Hedberg, N. & Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2): 274–307.

Harris, J. A. 2012. *Processing Perspectives*. PhD dissertation, University of Massachusetts Amherst.

Harris, J. A. 2014. *Shifting Viewpoints and Discourse Economy*. Poster presented at the 27th annual CUNY conference on human sentence processing, Ohio State University, USA.

Hegarty, M. & Waller, D. 2004. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence* 32(2): 175–191.

Hewitt, L. 1995. Anaphor in subjective contexts in narrative fiction. In *Deixis in Narrative: A Cognitive Science Perspective*, G. B. J. F. Duchan & L. Hewitt (eds), 325–339. Hillsdale NJ: Lawrence Erlbaum Associates.

Holler, A. & Irmen, L. 2007. Empirically assessing the effects of the right frontier constraint. In *Anaphora: Analysis, Algorithms and Applications. DAARC 2007. LNAI-Lecture Notes in Artificial Intelligence*, A. Branco (ed.), 15–27. Berlin: Springer.

Jaeger, T. F. 2008. Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4): 434–446.

Kaiser, E. & Cohen, A. 2012. *Free indirect discourse and perspective-taking*. Poster presented at the annual conference Architectures and Mechanisms for Language Processing (AMLaP), Riva del Garda, Italy. https://dl.dropboxusercontent.com/u/7489928/amlap2012.pdf

Kintsch, W. 1998. *Comprehension: A paradigm for cognition*. Cambridge, MA: CUP.

Kotovych, M., Dixon, P., Bortolussi, M. & Holden, M. 2011. Textual determinants of a component of literary identification. *Scientific Study of Literature* 1(2): 260–291.

László, J. 1986. Same story with different point of view. *SPIEL* 5: 1–22.

Long, D. L. 1994. The effects of pragmatics and discourse style on recognition memory for sentences. *Discourse Processes* 17(2): 213–234.

Maier, E. 2012. Switches between direct and indirect speech in ancient Greek. *Journal of Greek Linguistics* 12(1): 118–139.

Maier, E. 2014. Mixed quotation: The grammar of apparently transparent opacity. *Semantics and Pragmatics* 7: 1–67.

Nieuwenhuis, R., Pelzer, B. & te Grotenhuis, M. 2012. influence.me. Tools for Detecting Influential Data in Mixed Effects Models. Version 0.9.2.

Sanford, A. J., Moar, K. & Garrod, S. C. 1988. Proper names as controllers of discourse focus. *Language and Speech* 31(1): 43–56.

Schlenker, P. 2004. Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind & Language* 19(3): 279–304.

Sharvit, Y. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31(3): 353–395.

Simmons, J. P., Nelson, L. D. & Simonsohn, U. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* 22(11): 1359–1366.

Yao, B. & Scheepers, C. 2011. Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition* 121(3): 447–453.

What is a *Narration* – and why does it matter?

Sonja Zeman
University of Bamberg

Despite the increasing interest in textual structure and the specific use of lin-guistic means in narratives, the concept of *Narration* has so far not been an issue of theoretical concern within linguistics but has rather been used than defined or problematized. In this respect, the chapter takes a closer look at the concept of *Narration* itself. The starting point is the basic assumption that the analysis of Free Indirect Discourse (FID) – a phenomenon which seems to be restricted to *Narration* only – will offer crucial insights with respect to characteristics of the underlying narrative structure when read "against the grain". In this regard, the alignment of formal approaches of context shift and cognitive approaches of perspectivization leads to the conclusion that the narratological differentiation between narrator and character can be seen as a projection of the grammatical differentiation between speaker vs. observer, which is reflected on the different linguistic levels in a recursive manner. Consequently, I argue that the complex unfolding of different layers of discourse allows for poly-perspectival resp. me-tarepresentational effects on the textual macro-structure and hence constitutes a core characteristic of *Narration.* Against this background, the proposed model is not only able to align micro- and macro-structural accounts of *Narration* but also allows for a new perspective on the specific use of grammatical means in narrative discourse. In this respect, the paper argues that *Narration* does indeed matter for linguistics.

1. The mystery of *Narration*

> What are narratives? This is a tricky question […]
> (Gallagher & Hutto 2008: 30)

This chapter addresses a very basic question: What is a *Narration* actually? It is by no means a trivial question as it is intricately linked to the "mystery of *Narration*": while everyone intuitively seems to know what a narration is, no one knows ex-actly how to define it. Within the field of narratology, this elusiveness has led to an extensive discussion about the specific features of *Narration* and related concepts

such as 'narrativity', 'narrativehood', 'narrativeness' etc. (cf. e.g. Prince 2008). As a consequence, a multitude of definitions for *Narration* and narrativity has been proposed – however, we do not find a single definition everyone could agree on (cf. for a detailed overview Zeman 2016a).

In linguistics, by contrast, the mystery of *Narration* has led to quite the contrary: As the concept of *Narration* is commonly taken for granted within the linguistic tradition, it has never been a big theoretical issue. Although numerous studies investigate the behavior of grammatical entities in narrative texts, the status of *Narration* in everyday communication, and the acquisition of narrative tasks, among others, the very concept of *Narration* itself has not been seen as a major problem within the scientific debate and is, as a quite unproblematic presupposition, rather used than defined or discussed (cf. also Eckardt 2012: 17, footnote 8 with respect to the fact that a general theory of *Narration* is missing).

On the other hand, the relevance of narrative structure has been emphasized in many linguistic studies which show that the underlying discourse mode has an essential effect on – to name only a few – the usage of aspect, tense, and modal forms (cf. e.g. Dahl 1985; Fleischman 1990; Smith 2003, 2004), inference-drawing and implicatures (Caenepeel & Moens 1994; Asher & Lascarides 2003; Asher 2012), and structures of perspectivization (Carroll & Stutterheim 2007). Furthermore, there is a multitude of studies dedicated to the specific use of narrative devices such as Free Indirect Discourse (FID) and the Historical Present (HP) (cf. e.g. Banfield 1982; Fludernik 1992, 1993; Schlenker 2004; Sharvit 2008; Maier 2012; Eckardt 2014).

Hence, there are good reasons for assuming that there must be some difference between narrative and non-narrative discourse mode(s). But what exactly constitutes this difference? In order to answer this question, it seems reasonable to take linguistic devices as a starting point, which are considered to be particular narrative features of a text such as, for example, FID and HP. If it is correct that both phenomena do not appear in non-narrative contexts but are restricted to narrative discourse only, their linguistic analysis should lead us to new insights into the very nature of narrative discourse itself. These insights, in turn, should further be able to modify and refine previous accounts of the mysterious features of *Narration*. In order to pursue this line of reasoning, the chapter is organized as follows.

Section 2 gives an overview of conceptions of *Narration* in linguistics. Based on a comparison with narratological studies, I will argue that the common linguistic definition of *Narration* as 'a sequence of events' is not tenable to a general theory of *Narration* as it neglects the double-layered structure of narrative discourse as a crucial characteristic of narrative macro-structure. Section 3 defines the double-layered structure in linguistic terms. By taking formal and cognitive analyses of FID as a starting point (3.1), I will argue that the general principles of

context shift resp. viewpoint split are central mechanisms of narrative discourse. Interestingly, these mechanisms are not restricted to narrative discourse only (3.2). In order to account for the characteristics of narrative structure, I will take a closer look at the concept of the speaker (3.3) resp. narrator (3.4), which will finally lead to a differentiation between narrative and non-narrative discourse mode (3.5). I will argue that both modes are equally reliant on a double-layered structure that comprises the level of an illocutionary resp. 'narrative force' on the one hand and the level of proposition resp. the represented character/events on the other hand. However, unlike non-narrative discourse, where the two layers coincide by default, narrative discourse is constituted by a more complex configuration. This configuration is based on a distance between the illocutionary and locutionary level triggering an additional intensional layer of narratorship (i.e. 'narrative force') as an additional projection of the illocutionary level. In order to substantiate its generality, I discuss consequences of this assumption for the characteristics of *Narration* in spoken language (Section 4) and the matter of fictionality (Section 5). Finally, Section 6 will investigate the modal verb construction *sollte* + inf. in narrative discourse. I will argue that the proposed model is able to account for an alignment of the micro-structural usage of grammatical means and viewpoint effects on the macro-structure. In this respect, it will become obvious that the general aim of a theory of *Narration* seems indeed worth pursuing.

2.   What is a *Narration*?

> the humblest narrative is always more than a chronological series of events […]
> (Ricoeur 1980: 178)

2.1   Micro-structural conceptions of *Narration*: The sequence of events

As seen in the introduction, searching for the characteristics of narratives seems to be a rather tricky task as neither narratology nor linguistics has offered a definition of *Narration* commonly accepted. What constitutes the narrative impact of, for example, a fairy tale like *Little Red Riding Hood* has thus remained quite a mystery. As argued in Zeman (2016a), the trickiness of the question results from the fact that narrativity is not an absolute concept but refers to different phenomena from different perspectives. One could, for example, come to quite different conclusions with regard to the narrativity of *Little Red Riding Hood* by either investigating its linguistic macro-structure, i.e. the story as a whole (as it has been primarily the case in literary and narratological studies), or its micro-linguistic structure, which naturally has been the main focus in linguistics.

The distinction between micro- and macro-structure is evidently an over-simplification since different conceptualizations of *Narration* have been proposed both with regard to the macro-level (for a summary, cf. Prince 2003; Zeman 2016a) and the micro-level structure. This also means that there is no homogeneous conception of *Narration* on the micro-level as the term has been applied to narrative discourse relations, narrative discourse mode, as well as narrative context. Asher and Lascarides (2003), for example, argue that *Narration* constitutes one of several rhetorical relations that relate two "elementary discourse units", i.e. two propositions. The narrative relation holds if it "entails that the descriptive order of events matches their temporal order" (Asher & Lascarides 1993: 3). *Narration* in this sense is thus an abstract discourse relation which is typically inferred by the listener but can be indicated by linguistic devices such as discourse particles like *and then* and *next* which "monotonically yield *Narration*" (Asher & Lascarides 2003: 202). This is a rather restrictive conception of *Narration*, as only descriptions of events which match the "natural event-sequence" (Asher & Lascarides 2003: 200), i.e. the iconic succession of events, are seen as related by a narrative connection. In this sense, (1) would be a *Narration* (more precisely: a complex structure consisting of five distinct discourse units linked by the relation of *Narration*), while (1′), by contrast, would not:

(1) The wolf pressed the latch, * and the door opened. * He stepped inside, * went straight to the grandmother's bed, * and ate her up.

> [Grimm Brothers, Little Red Cap,
> < http://www.pitt.edu/~dash/grimm026.html > (11 August 2013)]

(1′) The door opened. The wolf had pressed the latch.

In (1), the discrete discourse units (indicated by asterisks) are constituents linked by narrative relations. The event described by the first proposition temporally precedes that of the second, and so on. The rhetorical relation in (1′), by contrast, is different. The second clause serves as an explanation of the first. According to Asher and Lascarides (2003: 7), the relation between the two propositions is not one of *Narration* but of *Explanation*. Within a framework of different *discourse relations*, such a definition seems convincing. However, by focusing on the *discourse structure*, one could come to different conclusions, as (1′) is, after all, part of a narrative context. Seen from this perspective, (1′) would also be narrative, as the narrative structure is not constituted by the fact that *all* relations between discourse units are narrative connections (as not all events in a narrative like *Little Red Riding Hood* are temporally sequenced). Hence, (1′) is both narrative and non-narrative at the same time: non-narrative because it is explicative in opposition to narrative, and narrative, because it is an explicative relation within a narrative – and hence part of a more general narrative discourse structure (including narrative and non-narrative parts).

But what, then, defines narrative *discourse mode*? According to Carroll (2001), it is the possession of narrative connections, which is "an essential feature of anything that we would want to call a narrative" (Carroll 2001: 22).[1] This is in line with the premise underlying most linguistic studies that it is the temporal order of events, which constitutes narrative discourse mode (cf. among many e.g. Fleischman 1990; Caenepeel & Moens 1994; Molendijk 1994; Smith 2003, 2004). However, narrative discourse mode obviously contains more than just narrative relations in the sense of Asher and Lascarides (2003). Furthermore, the sequence of events does not seem to be a sufficient criterion either, cf. (2):

(2)   Next comes the milk. Add in the milk and whisk until combined (it will get frothy). Finally comes the flour! Use a sifter and sift in the flour/cocoa mixture. Beat until the clumps are gone […].   [<http://www.activelifecooking.com/2009/02/choco-strawberry-clafouti/> (11 August 2013)]

Example (2) is clearly constituted by a representation of events in an iconic temporal order of events. Nonetheless, there is a clear intuition that (2) is not a *Narration* but an instruction for a cake recipe. The sequence of events is therefore not a sufficient criterion in order to define narrative discourse. For that reason, additional features are necessary to define narrativity. In some accounts, the specificity of the described events is seen as a necessary prerequisite to render their temporal sequence into a narrative (cf. e.g. Fleischman 1990: 79). Furthermore, narratives are seen as characterized by an anaphoric pattern where the events are related to each other like a temporal chain and constitute the reference point in form of a 'story now'. This implies that the temporal advancement of the events and the deictic displacement of the actual speaker's origo are seen as the relevant features of narrativity in these accounts (cf. e.g. Dahl 1985: 112; Wüest 1993: 232; Caenepeel & Moens 1994; Smith 2003, 2004). Moreover, it has also been argued that the sequenced events must be foregrounded (cf. e.g. Hopper 1979: 215), that the sequence of events implies that there must be more than two events (Carruthers 2005: 13), and, as already mentioned above, that the sequence of events is only narrative if the events iconically represent their "natural" sequence (cf. e.g. Labov 1972; Asher & Lascarides 2003).

Although this overview cannot be complete, it nevertheless demonstrates that linguistic definitions of *narrative discourse mode* display the pattern 'sequence of events + x', whereby the temporal order is seen as an absolute term, while x constitutes a variable. Insofar, it becomes obvious that the "priority of event structure"

---

1.   Note, however, that Carroll's (2001) conception of *Narrative Connection* cannot be equated with the discourse relation of *Narration* by Asher and Lascarides (2003) as Carroll (2001) argues for the fact that the *Narrative Connection* is constituted of more than temporal order.

underlies, as a presupposition, almost all linguistic conceptions of *Narration* (cf. also Fleischman 1990: 95). However, if one expands this view beyond discourse mode and refers to a wide concept of narrative context, it becomes questionable whether the sequence of events is a sufficient micro-linguistic feature. Consider the following example.

> (2′)   Next comes little red riding hood walking through the forest. Be careful, my dear child, and watch out for the wolf!

With respect to its linguistic structure, (2′) does not differ much from (2) above. Yet whereas (2) was intuitively classified as a non-narrative instruction, (2′) could be part of a narrative.[2] What is it that makes this structure a *Narration*? One thing seems clear: It must be something more than the sequence of events.

Altogether, it becomes evident that there is no reason to consider *Narration* as a rather unproblematic concept. On the one hand, we have seen that the term *Narration* is a relational concept as it is applied to different phenomena. Therefore, on the micro-level, we have to distinguish (at least) between *Narration* as a discourse relation and as a discourse mode. On the other hand, we have seen that the core principle of *Narration* remains unclear: though the main characteristic of *Narration* has been commonly seen within the representation of a sequence of events, the empirical facts suggest that this is neither a necessary nor sufficient criterion. Hence, the defining features are still unknown variables. In order to approach these unknown variables, let us see what accounts of *Narration* at the linguistic macro-level have to offer.

2.2   Macro-structural conceptions of *Narration*: The double-layered structure of discourse

Things become even more complicated if we take into account the narratological debates since there is no common agreement on what constitutes a Narration on the macro-structure at all. Is it a textual product? The producing act of storytelling? A macro-genre? A general mode of thinking? It is not the definition of particular artifacts that attend the main focus in narratology but the abstract notion of 'narrativity' covering the properties of narratives in general. Therefore, most narratologists assume that narrativity is best accounted for by a complex bundle of features. In such feature-based definitions the sequence of events is seen as one

---

2.   Indeed, even (2) could be part of a *Narration*, if, for example, the mother advises Little Red Riding Hood to bake a cake for her grandmother within the story.

of the most relevant characteristics, cf. the "minimal definition" by Genette (1982) (similar also Abbott 2008):[3]

> One will define narrative[4] without difficulty as the representation of an event or sequence of events.   (Genette 1982: 127)

Such a definition addresses one crucial aspect: *Narration* is not just seen as a sequence of events, but as the *representation* of it. In other words: every *Narration* is characterized by the distinction between the represented events and the manner of its representation, traditionally accounted for in terms of *story* vs. *discourse.* Leading back to the Russian Formalists (Šklovskij [1929²]1991; Tomaševskij [1925]1965; Propp 1928), this differentiation refers to the distinction between an abstract substance of the concrete events (i.e. *fabula*, corresponding to *story*) and their representation in discourse (*sjužet*, corresponding to *discourse*). With Chatman (1980: 19), one could say in a simplified way: "the story is the *what* in a narrative that is depicted, discourse the *how*". This distinction accounts for the rather trivial fact that a story can be represented in different manners and from different points of view. With respect to a definition of *Narration*, this differentiation entails one crucial implication. As there must be an organizing force behind every story, the concept of the narrator is seen as one of the most important characteristics of narrativity in many narratological accounts. This results in the essential distinction between two layers of discourse: the level of a presenting narrator and the embedded level of the represented events, cf. (3):

(3)   Double-layered structure of narrative discourse

---
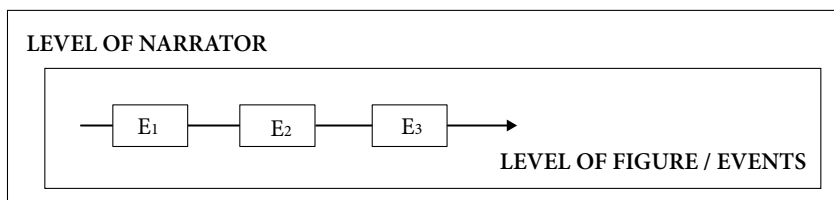
3.   In feature-based approaches (cf. e.g. Rimmon-Kenan 2006; Ryan 2007), narrativity is commonly seen as a gradual concept of prototypicality, whereby the family resemblance is accounted for by the amount of the narrative features of a particular narrative (for a critical review cf. Zeman 2016a).

4.   Genette (1982) refers here to 'narrative', not *Narration*, as the latter is in his tripartite account (i.e *histoire*, *récit*, and *narration*) defined as the narrative producing act within the narrative situation where it – actually or fictionally – takes place ("l'acte narratif producteur et, par extension, l'ensemble de la situation réelle ou fictive dans laquelle il prend place", Genette [1972] 2007: 15).

The differentiation between *story* and *discourse* in its several conceptions[5] can be seen as a presupposition underlying most narratological work. Consequently, the double-layered structure is commonly taken for granted in most definitions of narrativity. While linguists have mainly focused on the sequence of events on the story level, narratological definitions as the one given above rely on the premise that the narrator is one of the basic features of *Narration* since there is a narrator behind every story. Although this claim has been challenged by e.g. Banfield (1982), Fludernik (1993), and Patron (2009), who claim that there can be actually *Narrations* without a narrator in a rather abstract sense, "[f]ew would deny that narratives – or indeed any form of discourse – have their 'what' and their 'how,' whatever the terminological and conceptual mazes lurking beneath the surface of these apparently straightforward terms" (Pier 2003: 74). In addition, the distinction between *story* vs. *discourse* in this general sense is self-evidently not restricted to narratives only. Hence, it is questionable whether the double-layered structure could really serve as a distinctive feature of *Narration*. As the difference between the "set of structures in the represented world" (Titzmann 2003: 186) and the ways in which it is semiotically represented is an inherent property of all kind of utterances, this cannot be a characteristic that is specific for *Narration* only (cf. also Titzmann 2003). This leaves us with the question, what is the difference with respect to non-narrative discourse mode? In order to approach this question, the next section addresses the characteristics of double-layered structure from a linguistic point of view.

3.   The anatomy of the double-layered structure of *Narration*

> But nonetheless, the two levels can be distinguished in all classes of texts […].
> (Titzmann 2003: 195)

3.1   The case of Free Indirect Discourse (FID)

Although the narrative double-structure has not been an explicit feature within linguistic conceptions of *Narration*, it has been an issue in linguistic studies of narratives, particularly in accounts concerning 'reflector mode' resp. Free Indirect Discourse (FID). This mode is, as has already been pointed out, restricted to

---

5.   Cf. Pier (2003) for a critical overview regarding the different theoretical conceptions of these terms and their expansion to ternary resp. four-level approaches.

narrative discourse mode only.[6] FID is commonly considered to be a form of reported speech or thought which shares features of direct and indirect discourse. While pronouns and tenses are shifted to match the narrator's perspective, other linguistic devices such as indexicals, expressives, inverted questions, and interjections are preserved like quotations and are thus interpreted from the characters' perspective. Although the mechanisms underlying the characteristics of FID are far from being totally understood (cf. e.g. Schlenker 2004; Sharvit 2008; Maier 2012; Eckardt 2014 for discussion), it is seen as an uncontroversial fact that FID blends together two different viewpoints, cf. (4):

(4)   She could not be too soon alarmed, nor send for Perry too often. It was a pity, perhaps, that he had not come last night; for, though the child seemed well now – very well considering – it would probably have been better if Perry had seen it.   [Jane Austen, *Emma*, example taken from Nikiforidou 2012: 180]

In (4), the deictic elements *last night* and *now* do not reflect the perspective of the narrator, but have to be interpreted in relation to the 'story now', and hence attributed to the character's point of view. Likewise, modifying adverbials like *perhaps* and *probably*, and the parenthesis (*very well considering*) are bound by the viewpoint of the character. On the other hand, third-person-pronouns and the past tense imply a narrator's external third-person-stance. FID, hence, "behaves like a quotation whose 'grammatical skeleton' (= the tenses and pronouns) had been modified to match the perspective of the narrator" (Schlenker 2004: 285), while "*everything except pronouns and tenses*" (Schlenker 2004: 284; emphasis in original) is read from the character's perspective.[7] By contrast, Maier (2012: 3) analyzes FID as

---

6.   It is often argued that FID is not only restricted to narrative discourse mode, but, more precisely, to *fictional* resp. *literary* discourse mode. Hence, it would be questionable whether the characteristics of FID can reveal anything about narrativity per se, or whether they are in fact indicators of literality resp. fictionality (or a particular combination of these features). However, as has been evidenced by Fludernik (1993) with reference to earlier studies, FID is not exclusively restricted to neither literary nor fictional discourse but is also proven for oral and non-fictional narratives (cf. Fludernik 1993: 83f.). This will be discussed in detail in Sections 4 and 5. What can be said for now is that the article argues that the general principle of *Narration* exposed for FID is not restricted to literary fictionality, but also holds for non-literary, non-fictional discourse. Nevertheless, this section focuses firstly on fictional narratives – for the simple reason that in such context the principle of *Narration* can be exposed in a most transparent way.

7.   See however Maier (2012) who shows that "not all and not only pronouns and tenses are exempt from shifting to the protagonist's perspective" (Maier 2012: 27). In consequence, he rejects the thesis that the different shifting behavior of tenses and pronouns in contrast to other linguistic devices is the result of a dichotomic grammatical difference. Rather, the difference has to be seen

"essentially quotation with systematically punctured 'holes'" of unquoted pronouns and tenses. The mixed structure of FID hence reconciles the two perspectives of narrator and character, commonly accounted for in terms of 'double voice' resp. 'bivocal mode' (cf. e.g. Pascal 1977; Rauh 1978; Fludernik 1993; Schlenker 2004; Sharvit 2008; Eckardt 2012, 2014; Maier 2012; see also Vandelanotte 2009: 246–251 for a discussion).

The difference between these two possible perspectives has been accounted for in terms of various theories of context shift. In this respect, Eckardt (2012) distinguishes between External Contexts C vs. Internal Contexts c where C stands for narrator contexts and c represents the character contexts. In contrast to Internal Contexts, an External Context is defined as an assignment function where C and c are mapped to identical values. Hence, the difference is accounted for in the disintegration of External and Internal Contexts (cf. similar also Sharvit 2008). While in everyday life, external and internal contexts are, by default, the same, FID is characterized by a 'double context' (C,c) in the sense that FID has to be "evaluated relative to two utterance contexts $<C,c>$" (Eckardt 2012: 3).

Similarly, Schlenker (2004) draws a distinction between the Context of Thought (CoT, θ) vs. the Context of Utterance (CoU, υ). Whereas CoT (in FID: the character's perspective) is "the point at which a thought originates", CoU (in FID: the narrator's perspective) is "the point at which the thought is expressed" (Schlenker 2004: 279). While in everyday use, both contexts are as default the same, as the context of an intention and its expression is generally the same (cf. Schlenker 2004: 280), FID is characterized by the fact that it relies on the divergence between these two contexts. Hence, the key argument in Schlenker's (2004) account is that FID indicates a divergence of both contexts and thus provides evidence for a grammatical distinction that is not obvious in other discourse modes:

> The surprising fact, then, is that these literary styles [i.e. FID and HP; SZ] provide evidence for a grammatical distinction that has essentially no import in day-to-day life but seems to be hard-wired in language. (Schlenker 2004: 280)

Though this is not the only surprising fact of FID, it can be taken as a starting point for a model of *Narration*, relying on the basic assumption that accounts of FID allow for some basic implications with respect to the characteristics of narrative context when read against the grain. First, the structure of FID implies a 'double context'

---

as a "conventionalised, but ultimately pragmatic restriction on unquotation" (Maier 2012: 8). Similarly, Recanati (2010: 204ff.) argues that indexicals as *here* and *now* can be controlled both by the locutionary *and* the illocutionary context (i.e. Context of Thought vs. Context of Utterance in Schlenker's terms). Though this discussion is crucial with respect to FID, it does, however, not affect the general account of *Narration* proposed in the following.

which is the prerequisite for context shifts. Since FID is furthermore a stylistic device restricted to narratives only (cf. e.g. Banfield 1982, 1993) and prohibited in an "overall context" (cf. Eckardt 2012), it follows that the characteristics of FID as a particularly narrative device correspond to the characteristics of its narrative context and should hence reflect the configuration of narrative structures. In other words: if FID necessarily presupposes a double context and is, at the same time, as a particularly narrative device excluded from non-narrative passages, this would basically suggest that the narrative context provides the potential of context shifts by its double-layered structure. In consequence, this would lead to the rather trivial hypothesis that the core characteristic of a *Narration* is the availability of two contexts thus reflecting the double-layered structure of narrative discourse and the distinction between narrator and character as visualized in (3) above.

However, the next section will reveal that this cannot be the whole story but has to be refined in some crucial aspects. In this regard, the following section aims at looking behind the given intuition in refining the notion of narrative structure. In order to do so, it is not FID itself, which I will focus on in the following. Instead, previous analyses of FID are read against the grain in order to refine the insights with respect to the configuration of narrative *context*. In this respect, I will argue that FID is the surface phenomenon of a basic grammatical distinction between speaker and observer resp. evaluator, which is reflected on the different linguistic levels in a recursive manner. In this sense, I will argue that the distinction between the two contexts is, indeed, "hard-wired".

3.2   The *How* and *What* of narrative discourse

The last section has led to the suggestion that it is the double-layered structure of narrative discourse with its potential of context shift that configurates narrative discourse. The possibility of a context shift presupposes at least one more *possible* context besides the actual context. However, this does not say much about the quality of these two layers as it leaves open the questions of what distinguishes the context of a narrator and the context of a character in linguistic terms and to what extent such a distinction can account for narrative discourse structure. In order to address these two questions, let us have a closer look at FID. As pointed out above, the crucial characteristic of FID has been seen in the divergence of two different contexts, i.e. between an External Context (i.e. Context of Utterance) and an Internal Context (i.e. Context of Thought). This divergence provides the potential for a context shift. The dichotomy between 'external' and 'internal' corresponds to the narratological distinction in the tradition of Genette ([1972] 2007) between the narrator (*who speaks*) and the focalizer (*who perceives*). It can furthermore be linked

to the difference between *story* and *discourse*, as the speaker resp. narrator is seen as an element of external discourse while the internal thoughts of the character are attributed to the story level. In an abstract sense of the *how* and *what* of the story (or the manner of representation vs. the represented world), it is, however, self-evident that the double-layered structure cannot be a sufficient criterion to draw a distinction between narrative and non-narrative discourse mode. Not only narratives but also single utterances are naturally composed by, in the terms of Bally (1950: 36), the representation of the propositional content of an utterance (*dictum*) and the attitude of the modal subject (*modus*), the latter being the sentence's centerpiece related to the speaker's thought. Following the classical distinction between the content of a thought and its judgment (Frege 1979: 185), every sentence can hence be considered to be constituted by two different parts, namely its propositional content p and an attitudinal operator (Kiefer 1987: 75) resp. an illocutionary force (Searle 1975). In this sense, a sentence like (5) would presuppose two levels of discourse:

(5) This is Little Red Riding Hood's grandmother.
  → 'I hereby assert that p (This is little Red Riding Hood's grandmother)'

As illocutionary force is an indismissable part of every sentence (cf. e.g. Kiefer 1987: 75; Abraham 2012a: 62; Meibauer et al. 2013), the double-layered structure would hence be nothing special for narrative discourse but constitutive for every utterance. Though generally not marked explicitly within the linguistic structure but presupposed by the speaker, every illocutionary act is necessarily performed by a locutionary act (cf. Recanati 2010: 201). This double layered speech act configuration displays the potential for a split between the two discourse levels which becomes evident through linguistic means like epistemic modals, cf. (5′):

(5′) This must be Little Red Riding Hood's grandmother.

In (5′), the modal verb *must* indicates the degree of certainty that the speaker ascribes to the proposition p and hence his epistemic attitude towards the proposition. Epistemic markers like modal verbs are linguistic markers of the distance between the speaker's assessment and the content of the proposition (cf. in detail Leiss 2012 with respect to the double displacement of modal verbs). In this respect, epistemic modal verbs act like "condensed" propositional attitudes (cf. Pietrandrea 2005: 14):

(5″) This must be Little Red Riding Hood's grandmother.
  'I assume that p (This is Little Red Riding Hood's grandmother.)'

Propositional attitudes are in this respect "straightforward expressions" (Verhagen 2005: 98) indicating the double structure of representation just like non-factive

verbs of perception, cognition, or communication in the matrix clause indicate the perspective on the event while the embedded clause contains the perspectivized entity (cf. Verhagen 2005: 78). This structure has the potential of two different viewpoints, which can be identified by means of ambiguities with respect to the attribution of truth values, as illustrated in (6):

(6)   Little Red Riding Hood thinks that the wolf is her grandmother.
   → Yes, she does. (→ veridicality refers to the character's intensional state)
   → No, he is not. (→ veridicality refers to the truth value of the embedded proposition)

(6) contains two different contexts: The external context of the utterance and the internal context of Little Red Riding Hood's belief.[8] This becomes clear by the fact that *the wolf* can only be attributed to the speaker, as Little Red Riding Hood's belief system does not contain the fact that the creature lying in bed is the wolf. Hence, the sentence necessarily presupposes an external speaker who is not coded within the linguistic structure. In other words, there are two different belief systems involved. Furthermore, there is a hierarchical difference between these two viewpoints as the speaker's belief system is higher ranked insofar as the speaker is talking *about* the belief of Little Red Riding Hood. So what the example in the tenor of Frege shows is that propositional attitudes behave like narrative structure insofar as they provide two different contexts and hence the potential of a context shift.[9] In this sense, as has been already pointed out by Searle, speech acts and mental states are structurally analogous insofar as they share the dual structure, whereby the illocutionary force corresponds to the intentional state with an embedded propositional content (cf. also Recanati 2000: 128). This means that the hybrid structure is not excluded from non-narrative discourse. Consequently, the context shifting potential cannot be a sufficient criterion to account for the characteristics of narrative discourse. We can thus conclude with Titzmann (2003) that "the distinction between the two levels is not unique to narrative texts or texts with narrative structure" as it "can be distinguished in all classes of texts […]" (Titzmann 2003: 195). But how can we then account for the characteristics of narrative discourse? In the following subsection we address the question where these two contexts come about.

---

8.   It is discussed controversially whether verbs such as *believe* trigger a context shift or only a shift of circumstance (cf. e.g. Recanati 2000: 167f.). It is, however, taken as rather uncontroversial that such verbs are "world-creating" verbs insofar as they evoke a belief context distinct from the actual world (cf. McCawley 1978; Fauconnier [1985] 1994) and hence display the *potential* of a context shift.

9.   Self-evidently, the divergent contexts do not have to be linked to actual speakers – just as FID is not linked to a real narrator person and a real speaking character.

3.3   Is there a narrator at all?

In the previous section, it was argued that the double-layered structure linked to the distinction between speaker/narrator vs. character is equally crucial for the narrative and non-narrative discourse mode. However, we have already seen that FID is particularly characterized by the impression that there are only the thoughts of the character on the story level that are reflected as an unfiltered stream of consciousness, while there are no indications of a mediating instance. In consequence, Banfield (1982, 1993) and Fludernik (1993) have argued that FID is independent of a speaking subject and hence a kind of "non-communication". Hence, one can ask whether there is a speaker or narrator at all. Fludernik (1993: 443), reserving the term 'narrator' for a narrative voice constructed by linguistic devices such as first person pronouns, holds the view that pure reflector mode indeed displays *Narration* without a narrator, thereby supporting Banfield's (1982) claim that these (written) sentences are "unspeakable". However, Fludernik (1993) acknowledges that, "although […] the narrational process seems to refine itself out of existence by projecting a character's deictic centre of subjectivity, there is of course still the linguistic mediation by the narrative which appropriates the character's deixis" (Fludernik 1993: 442). In FID, this becomes clear by the personal pronouns and the past tense establishing a third-person stance, as past tenses like the simple past allow for the reconstruction of the narrator's position and, hence, for establishing a second layer of discourse, cf. (7):

(7)   a.   The wolf **presses** the latch.
            te = tr = ts
      b.   The wolf **pressed** the latch.
            (te = tr) < ts

Within the terms of Reichenbach (1947), the present tense is commonly accounted for by the coincidence of the time of speech (ts), the time of event (te), and the time of reference (tr), while the preterite is characterized by a shift of reference time and, hence, by a distance between reference time and time of utterance. In consequence, this deictic shift in the sense of Bühler's *Deixis am Phantasma* leads to the establishment of two different reference points: the original reference point linked to the time of utterance, and a second, displaced reference point in distance to the time of utterance. This leads to the effect that the time of event is doubly localized: the event is related, on the one hand, to the displaced reference point. At the same time, the relationship between the time of reference and the time of utterance is maintained, so that the time of event is also localized with respect to the original reference point, namely the time of utterance. Following Klein (1994: 140) and Johanson (2000: 34), the function of tenses goes beyond their temporal value insofar as they situate "the

perspective on the event rather than the event itself" (Johanson 2000: 34; emphasis in original). Due to this perspectival character that presupposes a point of view from which the focused events are described, tenses hence do not only localise an event in time but also inherently determine the speaking subject's origo (cf. Leiss 1992: 7).[10]

According to Leiss (2012), this "double displacement" leads to the effect that necessarily two different viewpoints arise. While the time of utterance is linked to the speaker's viewpoint, the displaced origo establishes a second viewpoint, bound to an observer resp. evaluator.[11] In other words: while the present tense displays the default case that the speaking and the observing subject coincide, the preterite mentally enables the speaker to establish two different viewpoints, cf. (7′)

(7′)   a.   The wolf **presses** the latch.
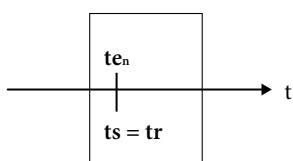
   b.   The wolf **pressed** the latch.
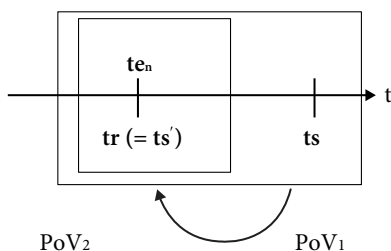
(PoV = point of view; ts = speech time; te = event time; tr = reference time; i.e. temporal point of perspective in the sense of Smith 2003: 100)

However, these two different viewpoints are not on the same level. As indicated by the boxes in (7b′), the original viewpoint includes the shifted one. Due to this relational character, the actual origo cannot be annulled, but is inherently a given. In other words, the simple past inherently always presupposes a speaker resp. a

---

10. Cf. Leiss (1992) for the general claim that the localizing of the observer's point of perspective lies at the foundation of grammatical functions. In this sense, the same principle holds also for the present tense: although the present tense involves no displacement, its relational structure nevertheless indicates the time of speech and hence allows for reconstructing the position of the speaking subject.

11. Needless to say, again neither "speaker" nor "observer" are actually real "persons".

narrator. This becomes obvious even in cases where the narrator seems to be effaced, cf. the examples of both tenses under embedding in (8).

(8) a.   Little Red Riding Hood believed that her grandmother **was** ill.
    b.   Little Red Riding Hood believed that her grandmother **is** ill.

The embedding of the present tense under an intensional verb like *believe* in (8b) triggers a different perspectival aspect than the embedding of a past tense in (8a).[12] While the simple past in the subordinate clause indicates that its content is attributed to the belief system of the subject of the sentence, i.e. Little Red Riding Hood, the present tense in (8b) is only valid when the time of the event is included within the time of utterance and the grandmother is still ill at the time of speaking. Hence, the content, which is asserted by the speaker, is not embedded within Little Red Riding Hood's belief system. In contrast, the simple past indicates subordination under the believe system of Little Red Riding Hood. At the same time, it further indicates that there is a report *about* a belief. Hence, as there must be some reporting instance, the speaker is inherently involved.

With respect to the tense usage in FID, this means that the speaker resp. narrator cannot completely be ruled out: even if there is no indication of a narrator persona, the narrator level is indicated in the deictic structure. According to Fludernik (1993), the tense usage in FID indicates the existence of an abstract narratorship in the sense of a "*structural* evocation of a sphere of (implicit) consciousness which provides the background for the plot experience" (Fludernik 1993: 212). Like Indirect Discourse, Free Indirect Discourse, though not syntactically embedded under a verbum dicendi, is subordinated to a "ruling narrating instance" (Fludernik 1993: 196).

With respect to narrative structure, the usage of tense has thus revealed two crucial aspects: first, the narrator level cannot be effaced but is a crucial prerequisite for narrative structure. Even in FID, where there are no indications of a narrator persona, there is an abstract level of narratorship as a structural property of narrative structure where the concrete narrator persona can or cannot be materialized (cf. in detail Zeman 2014).[13] Second, we have seen that grammatical means indicate the double layered structure thus displaying the potential of viewpoint shift

---

12. For the purupose of the present paper, the matter of tense in embedded sentences, as often referred to in terms of 'sequence of tense', is simplified, see Vandelanotte (2009) and Davidse and Vandelanotte (2011) for a comprehensive approach of speech and thought representation in English that takes also into account the perspectival viewpoint constellation in narrative and non-narrative discourse.

13. Fludernik (1993), in this respect, emphasizes that her view "does *not* imply that there is *always* a latent narrator" (Fludernik 1993: 453; emphasis in original).

by origo-displacement. In consequence, two differing viewpoints, which normally coincide, are present in narrative structures. With regard to tense, the present tense thus displays the "natural" default case, while with the past tense, the speaker is split into two: the speaker and a derived observer stance. In this respect, the relational structure of the preterite correlates with the split between the speaker (time of utterance) and the character level (time of event), and hence, the narratological distinction between the 'speaker' vs. 'the perceiver/focalizer'. The present tense corresponds to the default mode of reportative mode where speaker and focalizer resp. external and internal context coincide by default. Following the assumption of Bühler and Jakobson that the principle of shifting (origo-displacement resp. double displacement) has to be considered to be at the core of grammar (cf. Abraham 2012a,b; Leiss 1992, 2012), it is not surprising that the split between speaker and observer is, indeed, "hard-wired", and is also reflected on the discourse level in FID and propositional attitudes in a recursive manner. I will argue in the next section that this subject-split is also responsible for the structural discourse configuration that enables the context shift in *Narration*.

3.4   Who speaks?

So far, we have intuitively presupposed that the narrator is the speaker because he has been attributed to the external level of discourse. The narrator is the answer to Genette's question *who speaks?* In consequence, this leads to the assumption that the narrator is the same as the speaker of the story. Hence, the narratological claim that there is a narrator behind every story would mean nothing other than that behind every story, there is a speaker. Clearly, such a claim allows for no insight with respect to the specific nature of narrative structure, as it is an immaterial fact that non-narrative discourse mode also presupposes a speaker. So what is the narrator in linguistic terms then?

In order to address this question, let us have a look at FID again. According to linguistic accounts of context shift, the relevant criterion of FID has been seen in the fact that the two contexts involved in FID fall apart. In Schlenker (2004), this divergence is captured by the difference between the Context of Utterance (CoU) and the Context of Thought (CoT). Whereby CoU coincides with the actual context, CoT does not. However, there is more involved than just a context shift: as seen in the previous section, there are two different viewpoints linked to two different belief systems. Furthermore, we have seen with respect to Example (6) that the hierarchical difference between the two viewpoints is crucial, as the narrator/speaker is asserting something *about* the belief system of a third person on the embedded level. In this respect, Recanati (2010: 201) argues that it is not the distinction between the CoU and CoT which is relevant to account for the shift

but the difference between locutionary context (i.e. the context of utterance) and illocutionary context (i.e. the context of assertion). However, this move does not solve the entire problem either. Just like CoU and CuT, locutionary and illocutionary context are also seen, by default, as intricately linked together as a proposition can never be uttered without an illocution:

> An illocutionary act is taken to be performed *in* performing a locutionary act (Austin 1975), in such a way that there is a single context, and two possibilities.
>
> (Recanati 2010: 201)

Hence, the possibility that locutionary and illocutionary acts fall apart is not accounted for in traditional speech act theory (cf. Recanati 2010: 201),[14] whereas FID, where this principle is ruled out, requires that we drop the assumption that the context of assertion is necessarily linked to the context of utterance (cf. Recanati 2010: 202). Furthermore, in cases where the two contexts do not coincide, the hierarchical relationship of both levels becomes obvious, as the illocutionary act "is not actually performed, but is merely displayed, represented" (Recanati 2010: 202). In this respect, the narrative act has been seen as one of "pretense" (cf. e.g. Searle 1975; Recanati 2010):

> The objective features of the context of utterance are indeed 'given' and, to that extent, they cannot be shifted. But what the speaker can do is *pretend that the context is different from what it is*. If the pretense is mutually manifest, it will be part of what the speaker means that the sentence is uttered in a context different from the actual context c. In such a situation a context shift does occur: there are two contexts, the actual context c in which the sentence is produced, and the pretend context c' in which the utterance presents itself as being produced.
>
> (Recanati 2010: 193; emphasis in original)

In other words: the dual context situation is the result of a displaced intensionality of a speaker. Against the background of the principles of *Shifting* (Jakobson) resp. *Deixis am Phantasma* (Bühler), we can account for this mechanism in a more neutral way in terms of origo-displacement. The illocutionary level of the speaker is shifted and constitutes the basis for the narrator level so that there are two illocutionary forces resp. two internal contexts involved: the one of the actual speaker, and that of its shifted viewpoint, namely the narrator's context. So the narrator of a *Narration* is *not* simply the speaker but a projection of its displaced illocutionary force. The observations so far allow for a refinement of the configuration of narrative structure, which will be outlined in the next section.

---

14. Similarly, this premise underlies also the "Priority of Speaker" by Banfield (1982, 1993): she distinguishes between the E(xpression)'s SPEAKER, i.e. the speaking subject, and the SELF (i.e. its subject-of-consciousness), which are commonly coreferential and realized by the same morphological form, but actually fall apart in cases like FID.

3.5   Interim conclusion: the configuration of narrative discourse structure

In the previous sections, I have argued that neither the double-layered structure nor the principle of context shift as displayed in FID is a specific feature of narrative discourse only. Rather, it has been shown that both non-narrative and narrative discourse rely on a double-layered structure which renders a shift in viewpoint possible. Furthermore, the principle of a subject split between speaker and observer has been traced back to a more general principle underlying grammar in general, as accounted for in terms of *origo-displacement* (Bühler) and *shifting* (Jakobson).

However, *Narration* seems to be special in one point. Whereas the default mode for non-narrative discourse is that speaker and observer coincide, narrative structure is characterized by a distance between its speaker resp. its narrator, and, in consequence, an embedded discourse level. This configuration results from a distance between the illocutionary and locutionary subject. In this respect, the narrator is a displaced projection of the illocutionary subject of the actual speaker. In consequence, *Narration* by default displays a potential of context resp. subject shift, cf. (9):

(9)   The double-layered structure of narrative discourse – revised
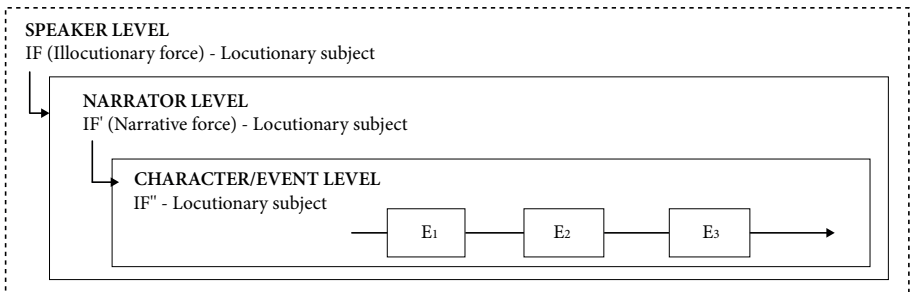
Hence, the discourse structure of *Narration* functions like mental state predicates insofar as propositional attitudes are also explicitly based on the distance between the content of an utterance and its assertion. This hierarchical difference between the different layers of narrative discourse is the prerequisite for poly-perspectival effects on the textual level such as FID which require the potential of alternate contexts resp. viewpoints. In this respect, FID is only a (possible) surface phenomenon of the underlying structure of discourse.[15]

Like the distinction between present tense and simple past is not seen to be a matter of different configurations but of the *displacement* between the

---

15. This is crucial to note as FID is a specific feature which is considered to develop not until the 19th century. Nevertheless, the anatomy of narrative structure is claimed to be valid also for historical stages of language (cf. in detail Zeman 2014).

temporal parameters, narrative discourse mode is also not generally different from non-narrative discourse mode but characterized by a more complex unfolding of different levels of discourse. As the principles of shifting resp. displacement are the same for non-narrative and narrative discourse, *Narration* can hence be seen as an extension of more general mechanisms (cf. also Zeman 2014).

But is it really *Narration* per se we are talking about? Remember that our considerations so far were based on the hypothesis that FID is a feature that is specific to the narrative discourse mode, and, in consequence, is able to lead us to the very principle of *Narration*. One might, however, object that our observations in Section 2 have also been restricted to *fictional literary* as well as *written narratives*. So can we really come to any conclusion with respect to *Narration* per se? Or have we in fact isolated a characteristic of *written literality* resp. *fictionality*? In order to discuss these questions, the next two sections will have a look at non-fictional resp. oral narratives – which will reveal that the outlined principle is not restricted to fictional resp. literary discourse, but is indeed a general principle of *Narration*.

4.   *Narration* in oral story telling

> In the oral narrative situation, narrative 'communicates' as much
> or as little as in written narratives – it tells a story.
> (Fludernik 1993: 442)

The starting point for approaching the nature of narrative structure has been the characteristics of FID as a linguistic device that is restricted to *Narration* only. However, FID is also a device that is considered to be restricted to *literary written Narration* only. Hence, is it really the structure of *Narration* that we have been talking about, or does this imply that the complex double-layered structure is restricted to *written* discourse and thus rather a feature of *literariness* resp. *literacy*? This consideration seems also particularly crucial within the context of the volume as it is linked to the question about what differences we have to expect between *Narration* in spoken vs. written language.

First of all, it is crucial to note that FID is not restricted to written *Narration* exclusively. Fludernik (1993) (with reference to other authors) offers wide evidence for the fact that FID is also proven in oral informal narratives (cf. Fludernik 1993: 83ff.). These examples suggest that the occurrence of FID is neither directly dependent on literacy nor literacy but rather displays an indirect affinity to written narratives.

This affinity between FID resp. *Narration* and written language has been thoroughly discussed in Banfield (1982, 1993), who has argued for a dichotomic difference between written and oral discourse. As already seen above, Banfield (1982,

1993) characterizes FID as "pure" *Narration* which is characterized by the absence of any mediating persona. According to her argument, such a constellation is only possible in non-conversational written contexts, as it is written language which "releases language from the communicative function" (Banfield 1993: 340). This leads her to suggest a strict bias between *Narration* vs. *Communication.* But is this opposition really clear cut? With respect to the observations laid out so far, two modifications seem to be crucial: first, it becomes evident by Banfield's argumentation that it is not the *medium* that triggers the differences between written and oral discourse in the sense of 'medial orality' (term according to Koch and Oesterreicher 1985), but the conversational situation, cf. Banfield (1993):

> More precisely, it is the dominance of the communicative function in speech which accounts for these differences in speech and writing.   (Banfield 1993: 340)

In this regard, the main difference between "oral" and "written" story telling lies particularly in the fact that oral *Narration* is necessarily bound to a conversational frame which implies that there is always a speaker who acts, at the same time, as a narrator. The question of a narrator's 'existence' – and his communicative function – is, according to Fludernik (1993) thus exclusively a matter in "written texts, with temporally *delayed* communication" (Fludernik 1993: 442; emphasis in original). The affinity between *Narration* and *written* texts hence becomes explicable as both *Narration* and written discourse display forms of displaced communication. However, this affinity is only an indirect effect linked to the communicative constellation and does not necessarily suggest a strict dichotomy between written vs. oral *Narration*, as self-evidently, *Narration* is not excluded from oral contexts.

Second, Banfield's conception of *Narration* as "non-communication" requires redefinition insofar as we have already seen that the mediating process is nevertheless coded inherently within the deictic structure of *Narration*, even if there is no narrator persona realized within the text. In consequence, this has led us to the differentiation between an abstract narratorship as a structural property of *Narration* and (possible) realizations of narrator personae (cf. Section 3.3). The necessity of this distinction becomes evident through a look at the default configuration in (oral) conversational settings. Let us assume a prototypical everyday oral *Narration* where a speaker tells a story that happened to him the day before. Say, for example, Little Red Riding Hood is narrating her story later to her mother. In this case, the speaker (Little Red Riding Hood) would be both the narrator of the story and its character. In consequence, the divergence between speaker, narrator and character would be seemingly effaced. However, though materialized in the same person, the different viewpoints linked to the different levels of discourse are not

necessarily the same: self-evidently, the temporal distance inevitably leads to different viewpoints, as Little Red Riding Hood's belief system at the time of the actual story and the time of storytelling are not the same. Coupled with that, the speaker in the role of the narrator takes an evaluative stance, so that the speaking subject is talking about himself as a depersonalized figure (as becomes obvious by the fact that a first person narrator can shift in third person pronouns when talking about himself, cf. in more detail Zeman 2014, 2016b). Hence, there is a general structural distance between the teller and the told that is independent of any difference from its medial realization resp. communicative situation.

Furthermore, there is also a divergence between the speaker and the narrator as the speaker constructs resp. refracts his story rather than gives a mimetic picture of the past events. This becomes clear by studies on everyday oral story-telling, cf. e.g. Günthner (2002). By examining embeddings of dialogic passages in oral everyday *Narration*, Günthner (2002) comes to the conclusion that the dialogues are not simple quotes but blend the perspective of different characters and the reporter's evaluation. Thereby, the difference between the animated figures and the evaluating voice of the reporter – in our terms: the difference between the story-level and the speaker/narrator – is indicated by metapragmatic devices such as prosodic features, voice quality, and code-switching. Thereby, the split between speaker and narrator becomes visible by the fact that the speaker does not only tell the events but acts as an narrator insofar as she "remodels the past text according to the situative communicative intention and imprints her perspective onto the reconstructed event" (Günthner 2002: 351). In this respect, Günthner (2002) discusses an example of first-person *Narration* where the reporter of a past dialogue presents herself as the protagonist and reconstructs her own utterance. This necessarily presupposes a split between the speaker-I and the narrator-I, leading to "a reconstruction of a reconstruction" (Günthner 2002: 354), as the I-speaker as I-narrator reconstructs the I-protagonist's reconstruction of the original speech (whereby it is irrelevant whether the "original speech" has indeed taken place in reality or not).

An even more obvious example of evaluative devices in narratives is furthermore the Historical Present (HP) which is commonly referred to as a characteristic device of particularly *oral* story telling (cf. among many e.g. Fleischman 1990; Fludernik 1993; Koch & Oesterreicher 2011: 170). With respect to the medial distinction between "oral" vs. "written", FID and HP hence display complementary distributions. Nevertheless, both narrative devices require the same textual structure as a prerequisite, namely the possible divergence of two different contexts. As Schlenker (2004) has shown, like FID also the HP is characterized by the fact that it is evaluated with respect to two different contexts, the Context of Thought (CoT) and the Context of Utterance (CoU):

> In Free Indirect Discourse, the actual context is the Context of Utterance, but the Context of Thought is presented as distinct, with the effect that someone else appears to be talking through the actual speaker's mouth ($c = \upsilon$, $c \neq \theta$). In narrations in the Historical Present the opposite pattern is found: the actual context is the Context of Thought, while the Context of Utterance is presented as having its time coordinate in the past ($c = \theta$, $c \neq \upsilon$). (Schlenker 2004: 299)

Under such an analysis, the HP constitutes the "mirror image" of FID since it displays "[t]he same distinction but the opposite pattern" (Schlenker 2004: 281). With respect to the conclusions to the properties of narrative structure, such a unifying account hence suggests that it is the potential divergence of contexts which constitutes the prerequisite for polyperspectival effects on the textual surface and, in consequence, a core characteristic of narrative structure. We can thus come to the conclusion that, though the complexity of poly-perspectivization might be most obvious in literary written texts, polyphonic strategies are also basic in non-literary oral narratives (cf. also Günthner 1999). Due to the different pragmatic setting of prototypical oral communication (face-to-face-situation, use of prosodic features, and multimodal devices as gesture and facial expressions), the narrative devices prototypically applied in spoken language may obviously be different. However, the mechanisms seem to be based on the same underlying narrative configuration. The distinction between an abstract narratorship and the textual materialization of narrator personae would hence suggest that the difference between oral and written discourse is linked to the different realizations of the narrator *personae*, while the structural properties remain the same. While this hypothesis would matter for further empirical studies, it is in line with the claim by Fludernik (1993) quoted above that FID, despite its obvious affinity to written discourse, is not categorically excluded from oral contexts.

Furthermore, recent studies on narration in sign languages give a clue that the potential of polyperspectivization that operates on the diverging levels of discourse is in no way restricted to written narratives only: in *Narration*, sign languages frequently use role shift as a specific kind of quotation. Interestingly, role shift is comparable to FID insofar as it combines features of direct and indirect speech, whereby a superordinated matrix verb can be omitted (cf. e.g. Herrmann & Steinbach 2012; Hübl 2013; Barberà & Quer this volume; Herrmann & Pendzich this volume). In this respect, both the oral-auditory modality of spoken languages as well as the visual-manual modality of sign languages display complex forms of polyperspectivization. Moreover, while perspectivization is materialized differently in the two different modalities, these differences do not seem to affect the structural configuration of *Narration* in general.

5.   What about fictionality?

> There is no textual property, syntactical or semantic,
> that will identify a text as a work of fiction.
> (Searle 1975: 325)

In the previous section, it turned out that the polyperspectival potential of *Narration* is not restricted to *written* discourse. But what about fictionality? FID in particular has been considered to be a marker of fictionality, as in "normal" life, we are not able to know the thoughts of persons other than ourselves. Reported thought in form of FID clearly seems to indicate that the discourse is fictional (cf. e.g. Hamburger 1957; Martínez & Scheffel 2003). This again raises the following two questions: (i) Is it really the principle of *Narration* per se we are talking about, or do we have to restrict our observations *fictional Narration* only? And if we do not restrict ourselves to *fictional Narration*, (ii) do we have to include the feature of fictionality in a definition of *Narration*? Both questions do not allow for an easy answer, as they are intricately linked to axiomatic premises about a theory of meaning. So, self-evidently, the following considerations cannot be more than a glimpse of the relationship between *Narration* and fictionality.

First of all, fictionality as well as literariness are concepts distinct from *Narration* and should not be equated. This is obvious in the fact that not all narratives are fictional (such as, for example, autobiographies, travel logs etc.) and not all fictional works are necessarily narratives. Fictionality and narrativity are hence two independent categories. The same holds for literariness since not all fictional narratives are literary, and "even if there were no such examples, it would still be a mistake [not to distinguish between the terms; SZ] because the concept of literature is a different concept from that of fiction" (Searle 1975: 320).

Taking the reference to the external world as the "crucial signpost of factual narration" (Martínez & Scheffel 2003: 232), *fictional Narrations* are characterized by their empty reference and hence do not allow for statements of veridicality. However, this is, at first place, a matter associated to the *fictionality* of fictional *Narrations*, not to the concept of *Narration* itself (cf. also Lamarque 2004: 398). With respect to narrative structure, it would be thus of no relevance whether the described story is referring to entities in the real external world or not. Considering the basic principle of *Narration*, a theory of fictionality (and literariness) is hence "in the first instance utterly irrelevant" because whether a story is fictive or non-fictive has "nothing whatsoever to do with the basic structures of narrative" (Titzmann 2003: 179).

So, if *Narration* and fiction are two different categories, we have to examine whether the general principle of the hybrid structure that we extracted from our observations of FID is a characteristic of fictionality or *Narration*. In this respect, it is relevant to point out that, just like FID is not restricted to oral narratives, it is

not restricted to fictional narratives either; cf. Example (10), given by Fludernik (1993) referring on Schiffrin (1981):

(10)   And w-
        So, the car stalled
        but we didn't ca-
        COULDn't call
        because we were supposed t'be out t'lunch
        and why were we HERE?
                [Schiffrin 1981: 47; quoted in Fludernik 1993: 84; emphasis in original]

(10) is evidently an example of non-fictional discourse, referring to a story of actual life experience. Hence, FID is not a feature of fictional discourse only, which in return means that the crucial feature we have extracted above – namely the hybrid structure comprising two possibly divergent contexts – is not necessarily restricted to fictionality. One could however claim (as actually an anonymous reviewer did) that the characteristic feature of FID lies within the fact that another person's thoughts are given without any indication where this insight comes from. This lack of information seems to be a special case insofar as it is only accepted by the readers/listeners in fictional contexts. However, this does not seem to be an inherent feature of FID, as it is ruled out in first-person examples like (10), where narrator and character are actually the same. Self-evidently, there is no need for the narrator to explicate the source of information of his former thoughts. Hence, the lack of information actually seems to affect the fictional aspect of FID, whereas, in contrast, the dual context situation is not exclusively linked to fictionality but a feature of *Narration* in general. Although narrator and character can be constituted by the same person, there is nevertheless a potential divergence of two different viewpoints, namely the actual speaker narrating the story and his former 'I' resp. 'we' which is bound to a different knowledge system. This holds irrespective of the story's fictionality. Remember the case of Little Red Riding Hood telling the story to her mother as discussed above in Section 3. The potential divergence of the two different viewpoints (e.g. the 'present-I' and the 'past-I') of an I-narrator is independent from the fact whether the discourse is fictional or not, i.e. whether Little Red Riding Hood or a neighbour's girl tells what happened to her – both stories would rely on the hybrid form of narrative discourse.

At a second glance, however, the problem of fictionality is not cast away so easily but hinges on the conceptualization of fictionality itself. Against the assumption of a gradual difference between 'reality' and 'fiction' in the sense of Hoorn (2012: 44), the concept of fictionality is not so much directly linked to the reference of an entity in the external world but to indirect observations of 'reality'. As fictional entities evade direct access and cannot directly be experienced within

the real world, their fictional status can only be experienced indirectly through a mode of communication (cf. Hoorn 2012: 45). In this sense, the conceptualization of fictionality is linked to the concept of mediacy, and, once again to the speaker.

In consequence, *Narration* by definition would display a lesser degree of certainty as it does not display a straightforward relation between the speaker and the world but a relation that is refracted by the level of an abstract narratorship. In this respect, the story level is perceived as a construction that emerges from textual perspectivization and cannot be directly validated against the external world. In order to "reach" the story level (may it be fictional or anchored in reality), we thus have to go along the epistemic path of different conceptualizers (cf. Langacker 2011). Such a premise would not be specific for fictional written narratives but for oral everyday narratives as well, as oral *Narrations* (as seen in Section 4) also display a construction of an experienced event (rather than a mimetic representation of the real world), cf. also Fludernik (1993):

> Mimesis in oral language would then be of the same quality as in fiction – not *imitation*, but invention and projection.   (Fludernik 1993: 426)

The difference between fiction and reality is hence a gradual one: while reportative discourse mode as a rule implies that the speaker stands behind his assertion, the narrator's assertion is, self-evidently, not the assertion of the speaker. Hence, there is a distance between the two levels before the story is being told, best seen in narrative configurations such as the unreliable narrator (cf. Zeman 2014). In this regard, aspects of fictionality, though not being a core feature of *Narration*, cannot entirely be excluded from observations of *Narration.*

What is crucial with respect to our main question concerning the structural characteristics of *Narration* is that the general principle of *Narration* – namely the hybrid discourse structure which sets the prerequisite for possible context shifts – is not exclusively bound to fictional narratives but to the characteristics of *Narration* per se. This has become clear by the fact that the fictional status of FID is dependent on the category of person. Examples of FID in first person show that the lack of a source of information as an indicator of fictionality is not an inherent feature of FID, but is bound to the different kinds of relationship between narrator and character. In contrast, fictional and non-fictional narratives are equally characterized by a complex discourse structure that comprises an additional level of narrative force and allows for polyperspectival effects. This is also supported by the fact that the Historical Present – as a narrative device, which is common in *non-fictional* contexts – blends two divergent contexts (cf. Section 3) and hence presupposes a hybrid structure of narrative discourse. Hence, while FID is in fact positioned within the interface of fictionality and *Narration*, the structural properties laid out above are not restricted to fictional narratives but display a core principle of *Narration* per se.

6. Why it matters

> […] the most important concept here is that of a **narrative context**.
>
> (Dahl 1985: 112; emphasis in original)

In Section 1, the main difference between linguistic and narratological definitions of *Narration* has been traced back to the fact that linguistic studies focus on the story level, while narratological studies rely on "mixed definitions" that combine the feature 'sequence of events' with the concept of mediacy (cf. Schmid 2003). The double-layered structure is hence an essential macro-structural feature of *Narration*. Do we "need" such a conception for a linguistic theory of *Narration*?

   With respect to accounts of FID, this seems to be rather obvious. Moreover, taking into account the hierarchical configuration of narrative discourse allows for a different perspective on other narrative devices, which have been primarily analyzed with respect to the story level. The empirical (non-fictional) example in (11) shows that the preterite form of the German modal verb *sollen* can be used as 'future of fate' in narrative contexts (resp. 'epic future', 'future in the past', 'futurum pro praeterito', see Zeman (2013a, 2014)):

(11) *Erst viele Jahre später **sollte** ich **begreifen**, welches Wissen mir von meiner Großmutter nebenbei vermittelt wurde.*
   'It was not until many years later that I **was to realize** [literally: '**should realize**'] what knowledge was incidentally passed on to me by my grandmother.'
   [<http://http://www.praxis-gradl.de/76-spagyrik.html> (25 February 2014)]

As (11) shows, the preterite form of the modal verb *sollen* in *Narration* is frequently used in order to give a foresight on events that will happen later with reference to the actual 'story now'. This prospective reading has caused troubles for the analysis of the modal verb construction as it does not seem to fit into the range of regular root resp. epistemic readings, cf. for comparison (12):

(12) *es **sollte** noch was im Kühlschrank **sein**. Ich ging in die Küche. Tatsache.*
   'There **should** still **be** something in the fridge. I went in the kitchen. Indeed.'
   → 'I suppose that there is something in the fridge. But I am not definitely sure.'
   [<http://www.joelle.de/topic/60812-der-anfang-vom-ende/page-2>]

Unlike the modal verb in (12), (11) does not signal a certain probability with respect to the evaluated proposition, whereby the time of event and the time of utterance coincide (cf. Leiss 2012: 55ff. for a detailed analysis of the epistemic use of *sollen*), but displays a high certainty for the event to come. In consequence, the use of *sollte* + inf. in cases like (11) has been considered to be a "special case" (Öhlschläger 1989: 176) and has commonly been accounted for in temporal terms (cf. e.g. Glas

1984: 101; Gloning 1991: 125; Duden 2009: §782, 820). In our terms: So far, the linguistic descriptions have addressed primarily the story level without taking into account the hierarchical structure of narrative discourse. The temporal description of the narrated event does, however, not fully account for its meaning as becomes clear in the following example without *sollen*:

(11′)   *Erst viele Jahre später* **begriff** *ich, welches Wissen mir von meiner Großmutter nebenbei vermittelt wurde.*
   'It was not until many years later that I **realized** what knowledge was incidentally passed on to me by my grandmother.'

As (11′) shows, posteriority with reference to the 'story now' would be indicated equally by the preterit without *sollen*. What is crucial, however, is the epistemic effect on the textual structure that is displayed by the use of the modal verb in (11), as it indicates a divergence between the narrator and the character level. While the narrator knows what will happen next, the protagonist on the story level, i.e. the sentence subject, is unaware of this. Hence, the difference between (11) and (11′) is not the temporal localization of the event within an epistemic effect, as the modal verb construction expresses a divergence between the knowledge systems of the narrator and the character. The complex unfolding of the different layers of discourse has the effect that the posterior event is not simply retold on the event level. Rather, the hierarchical level of narrative discourse induces that there is an assertive comment made by the narrator *about* the event.

This perspectival effect is in line with the deictic structure of epistemic modals as laid out in Section 3.2: modals act like condensed propositional attitudes insofar as they indicate a split between the locutionary and illocutionary subject. With respect to *sollen* + inf., there is even a further split between the speaker's knowledge system and the knowledge system of a third source, cf. Leiss (2012), for a thorough discussion. This deictic multiple displacement corresponds to a split between narrator and character level. Note, that (11) also illustrates the fact that the subject split is not bound to a morphological (not to mention actual) person.

Furthermore, the preterit form of the modal is responsible for the complex temporal localization. While the time of event is placed posterior with reference to the 'story now', the past tense indicates a displacement of the original speaker to a displaced reference time, i.e. to the level of the narrator. Consequently, the verbal situation as a whole is localized anterior with respect to the speech time resp. the speaker level and correlates to the hierarchical configuration of *Narration*, cf. (13):

(13) Temporal resp. epistemic structure of *sollte* + inf. as 'epic future'

(cf. Zeman 2014)
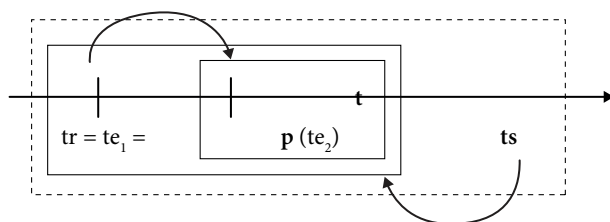
(ts = speech time; te_1 = event time (coded by modal verb); te_2 = event time (coded by infinitive); tr = reference time; i.e. temporal point of perspective in the sense of Smith 2003: 100; p = proposition)

We can thus conclude that an alignment of grammatical micro-structure and macro-structural point-of-view-phenomena depends on the double-layered configuration of *Narration* beyond the story level. The analysis of point-of-view-phenomena as epistemic effects of grammatical means on the textual surface suggests that it is in particular the investigation of ego-oriented linguistic means such as epistemic modal verbs, modal particles, and evidentials that will be able to offer new insights for the linguistic analysis of narrative structure, and, on the long run, for a general theory of *Narration*. In this respect, the observations made in this chapter can only represent the beginning of a longer discussion.[16] What has become more than obvious, though, is the fact that narrative structure indeed matters for linguistics and vice versa.

References

Abbott, H. P. 2008. *The Cambridge Introduction to Narrative*. Cambridge: CUP.

Abraham, W. 2012a. (Inter)subjectification or foreign consciousness/other's mind alignment as synchronic and diachronic concepts of change? Conceptualization and data fidelity. In *Covert patterns of modality*, W. Abraham & E. Leiss (eds), 24–78. Newcastle upon Tyne: Cambridge Scholars.

Abraham, W. 2012b. Illocutive force is speaker and information source concern. What type of syntax does the representation of speaker deixis require? Templates vs. derivational structure? In *Modality and Theory of Mind Elements across Languages*, W. Abraham & E. Leiss (eds), 67–108. Berlin: De Gruyter.

---

16. In this respect, it seems particularly promising to explore the compatibility of the present account with previous sentence-oriented definitions of *Narration* that consider shifting of reference point and 'distance' of the story world as crucial properties of narrative structure (cf. Dahl 1985; Caenepeel & Moens 1994; Smith 2003, 2004).

Asher, N. 2012. Implicatures in discourse. In *Discourse and Grammar. From Sentence Types to Lexical Categories*, G. Grewendorf & T. E. Zimmermann (eds), 11–42. Berlin: De Gruyter.

Asher, N. & Lascarides, A. 1993. A semantics and pragmatics for the pluperfect. In *Proceedings of the European Chapter of the Association for Computational Linguistics* (EACL 93), 250–259. <http://homepages.inf.ed.ac.uk/alex/pubs/pluperfect.html>

Asher, N. & Lascarides, A. 2003. *Logics of Conversation*. Cambridge: CUP.

Austin, J. L. 1975. *How to do things with words*. 2nd edn. Oxford: Clarendon Press.

Bally, C. 1950. *Linguistique générale et linguistique française*. Bern: Francke.

Banfield, A. 1982. *Unspeakable Sentences. Narration and Representation in the Language of Fiction*. Boston MA: Routledge & Kegan Paul.

Banfield, A. 1993. Where epistemology, style, and grammar meet literary history: The development of represented speech and thought. In *Reflexive Language. Reported Speech and Metapragmatics*, J. A. Lucy (ed.), 339–364. Cambridge: CUP.

Bühler, K. [1934] 1999. *Sprachtheorie. Die Darstellungsfunktion der Sprache*, reprint. Stuttgart: Lucius & Lucius.

Caenepeel, M. & Moens, M. 1994. Temporal structure and discourse structure. In *Tense and Aspect in Discourse* [Trends in Linguistics. Studies and Monographs 75], C. Vet & C. Vetters (eds), 5–20. Berlin: De Gruyter.

Carroll, M. & von Stutterheim, C. 2007. Durch die Grammatik fokussiert. *Zeitschrift für Literaturwissenschaft und Linguistik* 145: 35–60.

Carroll, N. 2001. On the narrative connection. In *New Perspectives on Narrative Perspective*, W. van Peer & S. B. Chatman (eds), 21–41. Albany NY: State University of New York Press.

Carruthers, J. 2005. Oral Narration in Modern French. *A Linguistic Analysis of Temporal Patterns* [Research Monographs in French Studies 19]. London: Legenda.

Chatman, S. B. 1980. *Story and Discourse: Narrative Structure in Fiction and Film*. Ithaca NY: Cornell University Press.

Dahl, Ö. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Davidse, K. & Vandelanotte L. 2012. Tense use in direct and indirect speech in English. *Journal of Pragmatics* 43: 236–250.

Duden. 2009. *Die Grammatik*, 8th edn. Mannheim: Dudenverlag.

Eckardt, R. 2012. Particles as speaker indexicals in free indirect discourse. In *Particles*, L. Hogeweg & E. McCready (eds.), 99–119. Special issue of Sprache und Datenverarbeitung.

Eckardt, R. 2014. *The Semantics of Free Indirect Discourse. How Texts Allow us to Mind-read and Eavesdrop*. Leiden: Brill.

Ehrlich, S. 1990. *Point of View: A Linguistic Analysis of Literary Style*. London: Routledge.

Fauconnier, G. [1985] 1994. *Mental Spaces. Aspect of Meaning Construction in Natural Language*, 2nd edn. Cambridge: CUP.

Fleischman, S. 1990. *Tense and Narrativity. From Medieval Performance to Modern Fiction*. London: Routledge.

Fludernik, M. 1992. The historical present tense in English literature. An oral pattern and its literary adaptation. *Language and Literature* 17: 77–107.

Fludernik, M. 1993. *The Fictions of Language and the Languages of Fiction: The Linguistic Representation of Speech and Consciousness*. London: Routledge.

Frege, G. 1979. *Posthumous Writings*, ed. by I. Baulbach, trans. by P. Long & R. White. Chicago IL: University of Chicago Press.

Gallagher, S. & Hutto, D. D. 2008. Understanding others through primary interaction and narrative practice. In *The Shared Mind. Perspectives on Intersubjectivity* [Converging Evidence in Language and Communication Research 12], J. Zlatev, T. P. Racine, C. Sinha & E. Itkonen (eds), 17–38. Amsterdam: John Benjamins.

Genette, G. [1972] 2007. *Discours du récit. Essai de méthode*. Paris: Éditions du Seuil.

Genette, G. 1982. Frontiers of narrative. In *Figures of Literary Discourse*, G. Genette (ed.), 127–144. Translation by A. Sheridan. New York NY: Columbia University Press.

Glas, R. 1984. ‚sollen' *im heutigen Deutsch. Bedeutung und Gebrauch in der Schriftsprache*. Tübingen: Narr.

Gloning, T. 1991. Bedeutung und zusammenhängendes sprachliches Handeln *Sollen* und das Vorausdeutungsspiel. In *Dialoganalyse III. Referate der 3. Arbeitstagung, Bologna 1990. Part 1*, S. Stati, E. Weigand & F. Hundsnurscher (eds), 123–134. Tübingen: Niemeyer.

Günthner, S. 1999. Polyphony and the 'layering of voices' in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics* 31: 685–708.

Günthner, S. 2002. Perspectivity in reported dialogues. In *Perspective and Perspectivation in Discourse* [Human Cognitive Processing 9], C. F. Graumann & W. Kallmeyer (eds), 347–374. Amsterdam: John Benjamins.

Hamburger, K. 1957. *Die Logik der Dichtung*. Stuttgart: Klett.

Herrmann, A. & Steinbach, M. 2012. Quotation in sign languages – A visible context shift. In *Quotatives: Cross-linguistic and Cross Disciplinary Perspectives* [Converging Evidence in Language and Communication Research 15], I. van Alphen & I. Buchstaller (eds), 203–228. Amsterdam: John Benjamins.

Hoorn, J. F. 2012. *Epistemics of the Virtual* [Linguistic Approaches to Literature 12]. Amsterdam: John Benjamins.

Hopper, P. J. 1979. Some observations on the typology of focus and aspect in narrative language. *Studies in Language* 3(1): 37–64.

Hübl, A. 2013. Role shift, indexicals, and beyond – New evidence from German Sign Language (DGS). In Proceedings of Texas Linguistics Society conference, L. C. Geer (ed.), 1–11. <http://www.leahgeer.com/tls13/huebl>

Johanson, L. 2000. Viewpoint operators in European languages. In *Tense and Aspect in the Languages of Europe*, Ö. Dahl (ed.), 27–187. Berlin: De Gruyter.

Kiefer, F. 1987. On defining modality. *Folia Linguistica* 21(1): 67–94.

Kindt, T. & Müller, H.-H. (eds). 2003. *What is Narratology? Questions and Answers Regarding the Status of a Theory*. Berlin: De Gruyter.

Klein, W. 1994. *Time in Language*. London: Routledge.

Koch, P. & Oesterreicher, W. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36(85): 15–43.

Koch, P. & Oesterreicher, W. 2011. *Gesprochene Sprache in der Romania*. Berlin: De Gruyter.

Labov, W. 1972. *Language in the Inner City. Studies in the Black English Vernacular*. Philadelphia PA: University of Pennsylvania Press.

Lamarque, P. 2004. On not expecting too much from narrative. *Mind and Language* 19: 393–408.

Langacker, R. W. 2011. The English present. Temporal coincidence vs. epistemic immediacy. In *Cognitive Approaches to Tense, Aspect, and Epistemic Modality* [Human Cognitive Processing 29], F. Brisard & A. Patard (eds), 45–86. Amsterdam: John Benjamins.

Leiss, E. 1992. *Die Verbalkategorien des Deutschen. Ein Beitrag zur Theorie der sprachlichen Kategorisierung*. Berlin: De Gruyter.

Leiss, E. 2012. Epistemicity, evidentiality, and Theory of Mind (ToM). In *Modality and Theory of Mind Elements across Languages*, W. Abraham & E. Leiss (eds), 39–65. Berlin: De Gruyter.

Maier, E. 2012. *Quotation and unquotation in free indirect discourse*. Ms, University of Groningen. <https://sites.google.com/site/emarmaier/publications>

Martínez, M. & Scheffel, M. 2003. Narratology and theory of fiction: Remarks on a complex relationship. In Kindt & Müller (eds), 221–237.

McCawley, J. D. 1978. World-creating predicates. *Versus* 19(20): 79–93.

Meibauer, J., Steinbach, M. & Altmann, H. 2013. Kontroversen in der Forschung zu Satztypen und Satzmodus. In *Satztypen des Deutschen*, J. Meibauer, M. Steinbach & H. Altmann (eds), 1–19. Berlin: De Gruyter.

Molendijk, A. 1994. Tense use and temporal orientation: The *passé simple* and the *imparfait* of French. In *Tense and Aspect in Discourse*, C. Vet & C. Vetters (eds), 21–48. Berlin: De Gruyter.

Nerbonne, J. 1986. Reference time and time in narration. *Linguistics and Philosophy* 9(2): 83–95.

Nikiforidou, K. 2012. The constructional underpinnings of viewpoint blends. The past + *now* in language and literature. In *Viewpoint in Language. A Multimodal Perspective*, B. Dancygier & E. Sweetser (eds), 177–197. Cambridge: CUP.

Öhlschläger, G. 1989. *Zur Syntax und Semantik der Modalverben des Deutschen*. Tübingen: Niemeyer.

Pascal, R. 1977. *The Dual Voice. Free Indirect Speech and its Functioning in the Nineteenth-century European Novel*. Manchester: Manchester University Press.

Patron, S. 2009. *Le narrateur: Introduction à la théorie narrative*. Paris: Armand Colin.

Patron, S. 2010. The death of the narrator and the interpretation of the novel. The example of Pedro Páramo by Juan Rulfo. *Journal of Literary Theory* 4(2): 253–272.

Pier, J. 2003. *On the semiotic parameters of narrative: A critique of story and discourse*. In Kindt & Müller (eds), 73–97.

Pietrandrea, P. 2005. *Epistemic Modality: Functional Properties and the Italian System* [Studies in Language Companion Series 74]. Amsterdam: John Benjamins.

Prince, G. 2003. Surveying narratology. In Kindt & Müller (eds), 1–16.

Prince, G. 2008. Narrativehood, narrativeness, narrativity, narratability. In *Theorizing Narrativity*, J. Pier, L. García & J. Ángel (eds), 19–27. Berlin: De Gruyter.

Propp, V. [1928] 1968. *Morphology of the Folktale*, 2nd edn. Austin TX: University of Texas Press.

Quasthoff, U. M. 1980. *Erzählen in Gesprächen. Linguistische Untersuchungen zu Strukturen und Funktionen am Beispiel einer Kommunikationsform des Alltags*. Tübingen: Narr.

Rauh, G. 1978. *Linguistische Beschreibung deiktischer Komplexität in narrativen Texten*. Tübingen: Narr.

Recanati, F. 2000. *Oratio obliqua, oratio recta: An essay on metarepresentation*. Cambridge, MA: MIT Press.

Recanati, F. 2007. *Perspectival Thought. A Plea for (Moderate) Relativism*. Oxford: OUP.

Recanati, F. 2010. *Truth-conditional Pragmatics*. Oxford: Clarendon Press.

Reichenbach, H. 1947. *Elements of Symbolic Logic*. New York NY: The Free Press.

Ricoeur, P. 1980. Narrative time. *Critical Inquiry* 7(1): 169–190.

Rimmon-Kenan, S. 2006. Concepts of narrative. In *COLLeGIUM: Studies across Disciplines in the Humanities and Social Sciences, Vol. 1: The Travelling Concept of Narrative*, M. Hyvärinen, A. Korhonen & J. Mykkänen (eds), 10–19. Helsinki: Helda. <http://www.helsinki.fi/collegium/journal/volumes/volume_1/>

Ryan, M.-L. 2007. Toward a definition of narrative. In *The Cambridge Companion to Narrative*, D. Herman (ed.), 22–36. Cambridge: CUP.

Schiffrin, D. 1981. Tense Variation in Narrative. *Language* 57(1): 45–62.

Schlenker, P. 2004. Context of thought and context of utterance. *Mind and Language* 19(3): 279–304.

Schmid, W. 2003. Narrativity and eventfulness. In Kindt & Müller (eds), 17–33.

Searle, J. R. 1975. The logical status of fictional discourse. *New Literary History* 6(2): 319–332.

Sharvit, Y. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31(3): 353–395.

Šklovskij, V. B. [1929²] 1991. *Theory of Prose*, ed. by B. Sher. Champaign IL: Dalkey Archive Press.

Smith, C. 2003. *Modes of Discourse. The Local Structure of Texts*. Cambridge: CUP.

Smith, C. 2004. Domain of tense. In *The Syntax of Time*, J. Guéron & J. Lecarme (eds), 597–619. Cambridge MA: The MIT Press.

Titzmann, M. 2003. The systematic place of narratology in literary theory and textual theory. In Kindt & Müller (eds), 175–204.

Tomaševskij, B. [1925] 1965. Thematics. In *Russian Formalist Criticism. Four Essays*, L. T. Lemon & M. J. Reis (eds), 61–89. Lincoln NE: University of Nebraska Press.

Vandelanotte, L. 2009. *Speech and Thought Representation in English. A Cognitive-functional Approach*. Berlin: De Gruyter.

Verhagen, A. 2005. *Constructions of Intersubjectivity. Discourse, Syntax, and Cognition*. Oxford: OUP.

Wüest, J. 1993. Textsortenabhängiger Tempusgebrauch. *Zeitschrift für französische Sprache und Literatur* 103(3): 231–243.

Zeman, S. 2013a. Zur Diachronie der Modalverben: *Sollen* zwischen Temporalität, Modalität und Evidentialität. In *Funktion(en) von Modalität*, W. Abraham & E. Leiss (eds), 335–366. Berlin: De Gruyter.

Zeman, S. 2013b. Vergangenheit als Gegenwart? Zur Diachronie des Historischen Präsens. In *Sprachwandel im Neuhochdeutschen*, P. M. Vogel (ed.), 236–256. Berlin: De Gruyter.

Zeman, S. 2014. (C)overt modality and its perspectival effects on the textual surface. In *Modes of Modality. Modality, Typology, and Universal Grammar* [Studies in Language Companion Series 149], W. Abraham & E. Leiss (eds), 457–484. Amsterdam: John Benjamins.

Zeman, S. 2016a. Introduction. In *Perspectives on Narrativity and Narrative Perspectivization* [Linguistic Approaches to Literature 21], N. Igl & S. Zeman (eds), 1–14. Amsterdam: John Benjamins.

Zeman, S. 2016b. Perspectivization as a link between narrative micro- and macro-structure. In *Perspectives on Narrativity and Narrative Perspectivization* [Linguistic Approaches to Literature 21], N. Igl & S. Zeman (eds), 15–42. Amsterdam: John Benjamins.

Reporting vs. pretending.
Degrees of identification in role play
and reported speech

Franziska Köder
University of Oslo

In both role play and reported speech, the actual speaker purports to present the speech of someone else. I analyse and compare the different perspectives that are involved in role play utterances and direct and indirect speech reports. I argue for a conceptual distinction between role play utterances and speech reports and discuss three criteria to distinguish them: (a) metalinguistic marking, (b) communicative intention, and (c) embodiment. Based on this analysis, I propose a hierarchy of identification with the other person, in which role play exhibits the highest degree of identification and indirect speech the lowest.

1. Introduction

At around two years, children begin to engage in role play and produce utterances as the character they impersonate (Harris 2000; Wolf, Rygh & Altshuler 1984). For instance, Wolf et al. (1984) describe a child at 24 months who adopts the perspective of a toy farmer and utters sentence (1) while placing the figure in a toy bath.

(1)   "Oh, no, sooooo hot, too hot. Ouch. Gotta put some cold in."   (age: 2;0)
(Wolf et al. 1984: 197)

At approximately the time of their first role play utterances, children also start producing direct speech reports (Ely & McCabe 1993; Nordqvist 2001a; Köder 2013). The reported utterances are in the beginning relatively simple, often consisting only of single words or onomatopoeia like animal sounds (see Example (2)).

(2)   "Muh" sagen die, ne?   (age: 2;1)[1]
'"Moo" they say, right?'

---

1.   All child language examples in this article, if not indicated otherwise, are taken from the CHILDES database (MacWhinney 2000).

During the third year, both role play utterances and speech reports become increasingly complex. Children produce now also indirect speech reports (Example (3)) (Köder 2013).

(3)   Mechthild sagt, dass du Pipi machen sollst.   (age: 2;7)
        'Mechthild says that you should do a wee-wee.'

Beside the fact that role play utterances and speech reports emerge developmentally at the same time, they also both involve two distinct perspectives. In role play (Example (1)), we can distinguish between a person's real-world identity (a child) and her identity in role play (a farmer). In speech reports (Example (3)), we can distinguish between the reporting speaker who is producing the speech report (a child) and the reported speaker whose speech is being presented (Mechthild). Thirdly, speech reports as well as specific aspects of role play (like role enactment, imaginary companions and role assignment) seem to be related to Theory of Mind in children (De Villiers & Pyers 2002; Lohmann & Tomasello 2003; Nielsen & Dissanayake 2000; Taylor & Carlson 1997). This could indicate that they are based on similar cognitive mechanisms like taking another person's perspective.

These similarities are the starting point for the following comparison between utterances in role play and direct and indirect speech reports. I start by distinguishing two notions of perspective, cognitive and linguistic perspective. I show that in contrast to indirect speech, in both role play utterances and direct speech the cognitive as well as the linguistic perspective is shifted to someone else. Despite this similarity, I argue for a conceptual distinction between role play utterances and speech reports. I suggest three criteria to distinguish them: (a) metalinguistic marking, (b) communicative intention, and (c) embodiment. Based on this analysis, I propose a hierarchy of different degrees of identification with the person whose speech is presented.

2.   Perspective

2.1   Two notions of perspective

Role play utterances and speech reports both involve two distinct perspectives: the perspective of the real-world identity and the perspective of the play identity in role play, and the perspective of the reporting speaker and the perspective of the reported speaker in speech reports. For the purpose of the following analysis, I distinguish two aspects of perspective: cognitive and linguistic perspective.[2] Cognitive

---

2.   The meaning of the terms "cognitive perspective" and "linguistic perspective" is similar to that of "focus location" and "viewpoint location" (Sanders & Redeker 1996).

perspective is the standpoint of the experiencing subject that is characterised by certain (linguistic) actions, perceptions, emotions and thoughts. Distinct from that is the linguistic perspective, that is the standpoint from which these actions and mental states are linguistically presented or – in other words – the deictic centre of an utterance. The deictic centre is the zero-point of the personal and spatio-temporal coordinate system (Bühler 1999 (1934)). The meaning of deictic expressions such as pronouns (*I*), spatial or temporal indexicals (*here*, *tomorrow*), evidentials (*probably*), evaluative or subjective expressions like epithets (*idiot*) needs to be determined relative to a specific point of orientation.

In the following, I will analyse in what way cognitive and linguistic perspectives are combined in role play as well as in direct and indirect speech reports.

2.2 Perspectives in role play

Role play involves two cognitive perspectives that correspond to a person's real-world identity and a person's identity in role play. Real self and play self can have different perceptions, world knowledge and evaluations of situations. Imagine for instance a child who plays a super hero with abilities such as reading people's thoughts or seeing in the dark.

Even though the cognitive perspective of the real self and the play self is conceptually distinct, people use their own knowledge schemata as basis for enacting a play character appropriately (Harris 2000). For instance, when children play doctor, they make use of their real-world knowledge of how such interactions are structured (e.g., investigation, diagnosis, prescription), what roles they involve (e.g., doctor, patient) and what typical behaviour – including linguistic behaviour – is associated with these roles (cf. Nelson & Gruendel 1986; Schank & Abelson 1977).

Cognitive perspective shifts in role play are connected to physical or bodily changes. The same cognitive perspective can be projected on different kinds of things in role play: one's own body (embodied role play), external objects (replica play) or nothing (imaginary companions) (Harris 2000).

In embodied role play, a person uses her own body to enact a character. Beside children, also actors in theatre or film engage in embodied role play. Characteristic of embodied role play is that the real self and the play self share and make use of the same body (Harris 2000; Sawyer 1996). To create a clear visual distinction between the two identities, the physical appearance of a person can be transformed by the use of guises or make-up.

In replica play, the second type of role play, external objects such as dolls, stuffed animals or toy figures are used as props. It is not uncommon in replica play that a person enacts multiple roles simultaneously. For instance, a child can pretend to be several toy figures and the narrator at the same time (Wolf et al. 1984).

The third type of role play, role play with imaginary companions, does not involve props that stand for the impersonated character. Imaginary companions are fictive persons or animals with certain stable properties, for instance "a 91 year-old man who is only 2 feet tall but can 'hit bears'" (Taylor & Carlson 1997: 446). Whereas embodied role play and replica play are common forms of play among all typically developing children, only some children create imaginary companions (Taylor 1999).

In all types of role play, not only the cognitive perspective of the play character is assumed but also his or her linguistic perspective. In Example (4), for instance, the sick play character and not the healthy child is the deictic centre of the utterance and therefore the referent of the pronoun *ich* 'I'.

(4) ich bin auch krank   (age: 3;4)
     'I am sick, too'

(5) möchten Sie auch Smarties?   (age: 3;6)
     'would you [formal] like Smarties as well?'

The shift of the deictic centre also changes social relations and requires people to adapt their language respectively. For instance, when playing shopping, the polite address *Sie* 'you' is more appropriate in German for the exchange between seller and customer than the informal *du* 'you' that children would use among each other and with their parents (see Example (5)). Children also adapt their pitch to the roles they enact. Nordqvist (2001b) gives the example of a three-year-old who lends the father doll a lower-pitched voice than the baby doll.

To sum up, in role play a cognitive perspective shift goes hand in hand with a linguistic perspective shift. This means that in accordance with the shift of the subject of experience also the deictic centre is shifted.

2.3   Perspectives in direct and indirect speech

Similar to role play, speech reports also involve two distinct cognitive perspectives: on the one hand, the cognitive perspective of the person who is producing the speech report, the so-called actual or reporting speaker; on the other hand, the cognitive perspective of the person to whom an utterance is attributed, the reported speaker. In a narrative, the narrator often assumes the role of the reporting speaker and presents the speech of the story protagonists (Bortolussi & Dixon 2003). It is also possible to report one's own speech. But even in this case, the distinction between reporting speaker and reported speaker is essential because the current self can be distinguished from past, future and hypothetical selves with respect to her personal and spatio-temporal coordinates as well as her mental states.

In both types of speech reports, direct speech (Example (6)) and indirect speech (Example (7)), the reporting speaker presents the utterance and the thereby expressed cognitive perspective of the reported speaker. However, the linguistic perspective from which the utterance is presented is different in direct and indirect speech. In direct speech, the deictic centre is shifted from the actual speaker (in the reporting clause) to the reported speaker Mary (in the quote). This means that the reported speech act is presented linguistically from Mary's internal perspective. Personal pronouns like *I* or spatial indexicals like *here* have to be evaluated with respect to Mary's personal and spatio-temporal coordinates at the context of utterance.

(6)   Mary said, "I l l like it here".

(7)   Mary stuttered that she liked it there.

In contrast to direct speech, the deictic centre is not shifted in indirect speech (Example (7)). Mary's utterance is presented from the external perspective of the reporting speaker. From his linguistic perspective Mary – who is not present when her speech is reported – is referred to with the third-person pronoun *she*. Since the reporting speaker's current location is different from the place where Mary produced her utterance, he uses the spatial indexical *there*. Unlike direct speech, where the actual speaker temporarily assumes both the reported speaker's cognitive and linguistic perspective, the situation is more intricate in indirect speech. The reported utterance expresses Mary's cognitive perspective, but it is linguistically presented from the reporting speaker's perspective. A consequence of this entanglement of perspectives are potential ambiguities between de re and de dicto readings of indirect speech reports (Coulmas 1986; Plank 1986).

The fact that an utterance in indirect speech is presented from the reporting speaker's linguistic perspective, constrains the linguistic material that can be used in the report. Unlike in direct speech, it is not possible to report expressive elements like *ouch* or to imitate a person's way of speaking (Banfield 1973; Coulmas 1985). If a reporting speaker wants to convey that Mary stuttered when she produced her utterance, he can demonstrate the stuttering in direct speech by imitating it (Example (6)), but he has to lexically describe it in indirect speech (Example (7)) (Clark & Gerrig 1990). For children, demonstrating an utterance of another person might be an easier and more effective way of communicating than describing it, especially when they lack the necessary descriptive vocabulary, like for instance the verb *stutter*. This could be one explanation for the finding that German and Dutch children use mostly direct speech between the ages of 2 and 4 (Köder 2013). Another reason for the predominance of direct speech could be that it is syntactically easier than indirect speech because it allows the integration of incomplete

sentences. In fact the majority of children's direct quotations do not contain finite verbs (Köder 2013).

Since direct speech can contain various expressive elements, it is often considered to be more vivid than indirect speech (Clark & Gerrig 1990; Coulmas 1985; Tannen 2007; Wade & Clark 1993; Wierzbicka 1974). This is supported by psycholinguistic and neurolinguistic studies that found an effect of reporting type on the perceptual simulations of readers. When reading sentences in direct speech as opposed to indirect speech, experimental participants were more likely to simulate the reported speaker's voice or talking speed (Yao, Belin & Scheepers 2011, 2012; Yao & Scheepers 2011; but see Eerland et al. 2013).

In the literature on speech reports, researchers tend to compare direct speech reports with drama or pretence. Wierzbicka (1974: 272) for instance writes that by quoting a person, the reporting speaker temporarily "assumes the role" of the reported speaker and "plays his part". Bortolussi and Dixon (2003) describe the effect of direct speech reports in narratives as follows:

> Readers view the characters speaking directly, as in play, and thus presumably experience a greater sense of immediacy or proximity to the characters.
>
> (Bortolussi & Dixon 2003: 202)

These characterizations suggest an intimate connection between direct speech reports and utterances in role play. This is in line with the previous analysis that shows that in both role play utterances and direct speech – but not in indirect speech – the cognitive as well as the linguistic perspective is shifted.

Despite these similarities, I argue that the distinction between role play utterances on the one hand and speech reports on the other hand should be maintained. In the following, I will suggest three main differences: (a) metalinguistic marking, (b) communicative intention, and (c) embodiment.

3. Differences between role play utterances and speech reports

3.1 Metalinguistic marking

In contrast to role play utterances, direct and indirect speech reports usually contain a metalinguistic marking (Bortolussi & Dixon 2003). It indicates that the presented utterance has to be ascribed to someone other than the actual speaker. A metalinguistic marking can have the form of a reporting clause (like in Example (8)), a prepositional phrase (Example (9)), an adverb (Example (10)), an evidential marker (like *-mış* in Turkish (Johanson 2003), Example (11)) or a subjunctive (Example (12)) from German (Fabricius-Hansen 2002).

(8) Mary said that the minister is sick.

(9) According to Mary, the minister is sick.

(10) Reportedly, the minister is sick.

(11) Bakan hastaymış.
     'The minister is reportedly sick.'[3]

(12) Der Minister sei krank.
     'The minister is supposedly sick.'

Metalinguistic markings in speech reports do not necessarily have to be provided by the reporting speakers themselves (Nordqvist 2001b). The following dialogue is taken from an interaction between a girl (age: 2;4) and her mother who jointly recall or make up – it is not clear whether the presented interaction has in fact taken place – a conversation between the child and doctor Maiburg.

Dialogue 1

1	Child:	Doktor Maiburg
		'Doctor Maiburg'
2	Mother:	ähä und was sagt der zu dir?
		'ähä and what does he say to you?'
3	Child:	ausziehn
		'undress'
4	Mother:	ja und was sagst du da?
		'yes and what do you say then?'
5	Child:	ja nein!
		'well no!'

In this dialogue, the child's utterances *ausziehn* 'undress' (line 3) and *ja nein* 'well no' (line 5) do not contain a metalinguistic marking. Nevertheless I argue that they are speech reports and not role play utterances because the metalinguistic marking is provided by the mother in the preceding linguistic context. The questions *und was sagt der zu dir?* 'and what does he say to you?' (line 2) and *und was sagst du da?* 'and what do you say then?' (line 4) already indicate whose speech will be reported. In the corresponding answer, the announced speech report can then be given without the necessity of an overt metalinguistic marking.

Speech reports have in contrast to utterances in role play two levels of content: an object-representation and a "'meta' part, whereby the object-representation is referred to as entity on its own right and situated in the order of things" (Recanati 2000: xii). In the case of speech reports, object- and meta representation are both

---

3. The Turkish example is taken from Johanson (2003).

utterances (Wilson 2000). There are clear differences in content and truth conditions between a sentence with or without a metalinguistic marking like *Mary said*. It can for instance be true that Mary said that the minister is sick even though the minister is in fact in the best of health.

The metalinguistic marking is presented from the cognitive as well as the linguistic perspective of the actual speaker, i.e., from a perspective external to that of the reported speaker. The effect of a metalinguistic marking becomes clearer when we imagine an actor who is playing Hamlet but instead of just uttering his lines, he always starts his sentences with the metalinguistic marking *Hamlet says*. The reporting clause introduces an external perspective that is combined with the internal perspective of Hamlet (Recanati 2000). It is an empirical question how an audience would react to such a change. My hypothesis is that the presence of a metalinguistic marking would destroy the dramatic illusion that it is Hamlet who is speaking. This could lead to a distance ('Verfremdungseffekt', Brecht) between the audience and the character in the play and decrease the degree of identification or empathy with Hamlet.

In role play, a person's real-world identity is also still latently available. Usually both the role-playing person and her interaction partners or audience share the knowledge that a given (linguistic) act is merely an act of pretence, i.e., a "non-serious action" (Goffman 1974). The crucial difference between utterances in role play and direct speech is that the external or real-world perspective is not linguistically represented and therefore less salient. Within the scope of the play world, only a person's play selves are available correlates for referring expressions. The real-world self of a person recedes completely into the background for the time of pretence. It can of course be re-activated at any time. But shifting the deictic centre back to the real-world self is usually connected with stepping out of the play frame.

In this section, I have suggested that speech reports and utterances in role play can be distinguished by the presence or absence of an overt or elliptic metalinguistic marking. But this criterion must be treated with caution in view of children's development. On the one hand, children's early speech reports often do not contain metalinguistic markings. On the other hand, play activities also include meta-communicative utterances like "let's pretend that I am a doctor" that exhibit features similar to direct speech reports.

Hickman (1993) studied children's speech reports in narratives. She found that 4-year-old English-speaking children, who had to re-tell a previously heard dialogue, often did not mark explicit boundaries between the narrator's perspective and the reported speaker's perspective. Instead of reporting speech, they re-enacted it like in role play without a meta-linguistic marking. This makes it hard for listeners to understand whose speech the child presents. Nordqvist (2001b) found a similar tendency in the frog story narratives of 3-year-old Swedish children. Nordqvist

(2001b) calls those un-embedded speech ascriptions in narratives "free direct speech" – the same term that she also uses when children lend their voices to dolls in the context of replica play. What these cases have indeed in common is that the speech of someone else is presented without a metalinguistic marking. Despite this similarity in form, there is an essential difference. In the case of role play utterances, the non-embedded, "free" form is already the target form. However, in the frog story narratives "free direct speech" means either an elliptic speech report or a proto-speech report lacking a metalinguistic marking. It would be desirable that these conceptual differences are reflected in the terminology.

Beside speech reports that look like role enactments, there is also a special class of utterances connected to role play that has similarities with direct speech. These so-called meta-communicative utterances are located on a scale in-between the real world and a play world (Griffin 1984). Their function is to set up a play frame with its specific premise system that differs from reality (Bateson 1973). Meta-communicative utterances are used to draw boundaries between the play-world and reality, to transform situations, persons, things and actions and to plan the plot (Andresen 2005). In (13), for instance, a child establishes who is going to play a certain toy figure.

(13)   Ich $_{real\ self}$ bin der [/] der Vater von mir $_{play\ self\ (toy)}$.   (age: 2;7)
       'I $_{real\ self}$ am the [/] the father of me $_{play\ self\ (toy)}$.'

An interesting property of utterances like (13) is that referring expressions have to be evaluated with respect to two different deictic centres, comparable to direct speech. The personal pronoun *ich* 'I' refers to the child's real-world identity, the pronoun *mir* 'me' by contrast refers to a play identity of the child, a toy figure. The function of the complete sentence is to determine the identity of yet another play self of the child: a toy figure that is the father of her other toy figure.[4] Unlike direct speech, there are no graphic (e.g., quotation marks), gestural (e.g., air quotes) or para-verbal features (e.g., pitch) (Couper-Kuhlen 1998; Klewitz & Couper-Kuhlen 1999) that indicate the beginning or end of a deictic shift.

With these two caveats in mind, metalinguistic markings are a criterion to distinguish direct and indirect speech reports from role play utterances.

---

4.   The identification of a person with different selves in different worlds has similarities with dream reports like "I dreamt that I $_{real\ self}$ was Brigitte Bardot and that I $_{dream\ self\ (Bardot)}$ kissed me $_{real\ self}$" (Lakoff 1970; Percus & Sauerland 2003).

3.2 Communicative intention

The second difference between role play utterances and speech reports are the communicative intentions associated with *pretending* versus *reporting*. In role play, a person *pretends* to be another person and wants her play partners or audience (in e.g., theatre or film) to share this fiction. In contrast to that, in direct and indirect speech the reporting speaker's intention is to *report* an utterance. In direct speech, the reporting speaker adopts temporarily the cognitive and linguistic perspective of the reported speaker in order to demonstrate to her addressees what it would be like to listen to the reported speaker directly (Clark & Gerrig 1990). But she does not change her identity or pretend to be the person whose speech she is presenting. This difference in communicative intention is reflected in the set of felicitous responses to a role play utterance as opposed to a speech report.

A test is to check whether it is possible to address a person with the shifted identity after or during she produces a role play utterance or a speech report. For instance, if Mary is impersonating a queen in role play, it is possible to address her as "Your Majesty" (a). By contrast, if she reports directly what the queen says (b), such an address is infelicitous, even if Mary is interrupted in the middle of the quote when the queen is the deictic centre. Example (c) shows that addressing a person with the shifted identity is even more infelicitous in indirect speech than in direct speech.

Test: Addressing a person with shifted identity

- (a) *Mary (pretending to be a queen):* I would like to have tea. (role play utterance)
       Peter: At your service, your Majesty.
- (b) *Mary:*   The queen says, "I would like to have tea".   (direct speech)
       ??Peter: At your service, your Majesty.
- (c) *Mary:*   The queen says that she would like to have tea.   (indirect speech)
       *Peter: At your service, your Majesty.

These differences in felicity correspond to a hierarchy of different degrees of identification with the shifted identity, i.e., the queen in this example. I suggest that role play utterances exhibit the highest degree of identification and indirect speech reports the lowest.[5]

---

5. See Sanders & Redeker (1996) for differences in perspectivization between direct and indirect speech.

3.3   Embodiment

A third possible difference between role play utterances and direct speech reports is that in role play a person does not just present the speech of someone else but also *embodies* that person. This means that she also demonstrates motoric aspects like for instance someone's gestures or walking style.

In the framework of embodied cognition, Noice and Noice (2001) and Scott, Harris and Rothe (2001) show that bodily movement has a positive effect on the mental representation of a character's speech. Adult participants remembered scripts best when they enacted a character's bodily movements while reading aloud the character's text. Pellegrini and Galda (1982) also found an effect of role play on children's story recall. 5- to 7-year-olds listened to a story and were then assigned to one of three groups: they either reenacted the previously heard story, discussed it or drew pictures about it. The results indicate that children's story recall, their story comprehension, and their evaluation of character's actions were best in the reenactment condition.

At issue is whether the effect of children's role play on their story representation is actually due to embodiment. It could also be the case that enacting a character's role leads to a deeper identification with a character than discussing his or her actions from an external perspective or producing drawings that only allow the presentation of visible (but not mental) content. The effect of embodiment can be investigated by contrasting embodied role play with other types of play such as replica play or role play with imaginary companions. Marbach and Yawkey (1980) compared the influence of embodied role play and replica play on story recall. 5-year-old children listened to an action-oriented story about a bear and had to either show the bear's actions with their own body or with a hand puppet. Embodied role play led to significantly better memory of the story than replica play, supporting the embodiment hypothesis. Marbach and Yawkey (1980: 262) explain this result with "greater direct, immediate, and apparently more meaningful involvement" when using one's own body rather than a hand puppet as medium for pretend actions. An open empirical question is whether embodied role play also leads to a higher degree of identification with a story character than replica play or role play with imaginary companions.

Since not all types of role play – only embodied role play but not replica play or imaginary companions – involve embodiment, it is not suited as a criterion to distinguish role play utterances from speech reports. Furthermore, also speech reports can be linked to embodied elements. In oral face-to-face communication, a reporting speaker can also demonstrate the reported speaker's facial expressions, gestures and other bodily movements. However, people are probably more likely to

use embodied features in (embodied) role play than in speech reports and more so in direct than in indirect speech. This means that embodiment is not a distinguishing criterion but rather a gradual feature decreasing from (embodied) role play to direct speech and indirect speech.

In addition to spoken languages, it would be interesting to compare the difference between role play and speech reports in sign languages. German Sign Language (DGS) has, for instance, a type of speech report called "role shift" that combines features of both direct and indirect speech (Herrmann & Steinbach 2012). Role shift is signalled by eye gaze, head position and body lean and can also be accompanied by demonstrations of the reported speaker's facial expressions (Herrmann & Steinbach 2012). Since sign languages standardly involve the visual-manual modality, they are likely to be more embodied than spoken languages. But it is an open question whether this kind of embodiment that is based on conventionalised signs gives rise to similar cognitive effects as embodied role play.

4.   Hierarchy of identification

In this paper, I have outlined the similarities and differences between speech reports and role play utterances. In direct speech as well as in role play, an utterance is presented from another person's cognitive and linguistic perspective. This is unlike indirect speech, where the reporting speaker remains consistently the deictic centre and only the cognitive perspective of the reported speaker is presented. Nevertheless, there are clear differences between speech reports and role play utterances.

First, speech reports usually contain an overt or elliptic metalinguistic marking like *Mary said* (see Table 1). This metalinguistic marking is presented from the cognitive and linguistic perspective of the reporting speaker. The fact that speech reports have two layers of content – the metalinguistic marking and the reported utterance – might complicate the comprehension and production of these constructions. This could be an explanation for the finding that younger children often fall back on the presumably easier enactment mode during story-telling (Hickmann 1993; Nordqvist 2001b).

Second, the communicative intentions associated with speech reports and role play utterances differ. Only in role play, a speaker is actually *pretending* to be another person (see Table 1). This is why it is felicitous to address the speaker with the shifted identity after or during a role play utterance in contrast to a direct or indirect speech report. A third candidate to distinguish role play utterances and speech reports was embodiment. I suggested a scale of different degrees of

Table 1. Comparison between role play utterances, direct speech and indirect speech

	Role play utterance	Direct speech	Indirect speech
Example	*I am sick.*	*Mary said, "I am sick."*	*Mary said that she is sick.*
Metalinguistic marking	–	+	+
Communicative intention	Pretending	Reporting	Reporting
Embodiment	***	**	*

embodiment decreasing from (embodied) role play to direct speech and indirect speech (see Table 1).

Based on the previous analysis, I propose a hierarchy of different degrees of identification with the person whose speech is presented. Role play utterances exhibit the highest degree of identification. In role play, a person pretends to be someone else and produces utterances from that person's cognitive and linguistic perspective. In direct speech, the cognitive and linguistic perspective is shifted to the reported speaker as well. However, the presence of the metalinguistic marking introduces an external perspective which reduces the degree of identification with the reported speaker. Indirect speech exhibits the lowest degree of identification. Only the content of the reported speaker's words is presented but from the reporting speaker's linguistic perspective. At issue is whether this hierarchy of identification has also psychological reality. It could be experimentally tested whether people's empathy or evaluation of a character is influenced by the type of utterance (role play utterance vs. direct speech report vs. indirect speech report).

Acknowledgements

References

Andresen, H. 2005. Role play and language development in the preschool years. *Culture & Psychology* 11(4): 387–414.

Banfield, A. 1973. Narrative style and the grammar of direct and indirect speech. *Foundations of Language* 10(1): 1–39.

Bateson, G. 1973. A theory of play and fantasy. In *Steps to an Ecology of Mind. Collected Essays in Anthropology, Psychiatry, Evolution and Epistemology*, G. Bateson (ed.), 150–172. London: Granada.

Bortolussi, M. & Dixon, P. 2003. *Psychonarratology. Foundations for the Empirical Study of Literary Response*. Cambridge: CUP.

Bühler, K. 1999 (1934). *Sprachtheorie. Die Darstellungsfunktion der Sprache. Mit einem Geleitwort von Friedrich Kainz*. Stuttgart: Lucius & Lucius.

Clark, H. H. & Gerrig, R. J. 1990. Quotations as demonstrations. *Language* 66(4): 764–805.

Coulmas, F. 1985. Direct and indirect speech: General problems and problems of Japanese. *Journal of Pragmatics* 9: 41–63.

Coulmas, F. 1986. Reported speech: Some general issues. In *Direct and Indirect Speech*, F. Coulmas (ed.), 1–28. Berlin: Mouton de Gruyter.

Couper-Kuhlen, E. 1998. Coherent voicing. On prosody in conversational reported speech. *Interaction and Linguistic Structures* 1: 1–28.

De Villiers, J. G. & Pyers, J. E. 2002. Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development* 17: 1037–1060.

Eerland, A., Engelen, J. A. A. & Zwaan, R. A. 2013. The influence of direct and indirect speech on mental representations. *PLOS ONE* 8(6): 1–9.

Ely, R. & McCabe, A. 1993. Remembered voices. *Journal of Child Language*: 20(3): 671–696.

Fabricius-Hansen, C. 2002. Nicht-direktes Referat im Deutschen. Typologie und Abgrenzungsprobleme. In *Modus, Modalverben, Modalpartikeln*, C. Fabricius-Hansen, O. Leirbukt & O. Letnes (eds), 7–29. Trier: Wissenschaftlicher Verlag.

Goffman, E. 1974. *Frame Analysis. An Essay on the Organization of Experience*. Cambridge MA: Harvard University Press.

Griffin, H. 1984. The coordination of meaning in the creation of a shared make-believe reality. In *Symbolic Play. The Development of Social Understanding*, I. Bretherton (ed.), 73–100. Orlando FL: Academic Press.

Harris, P. L. 2000. *The Work of the Imagination*. Oxford: Blackwell.

Herrmann, A. & Steinbach, M. 2012. Quotation in sign languages. A visible context shift. In, *Quotatives. Cross-linguistic and Cross-disciplinary Perspectives* [Converging Evidence in Language and Communication Research 15], I. Buchstaller & I. Van Alphen (eds) 203–228. Amsterdam: John Benjamins.

Hickmann, M. 1993. The boundaries of reported speech in narrative discourse: Some developmental aspects. In *Reflexive Language: Reported Speech and Metapragmatics*, J. A. Lucy (ed.), 63–90. Cambridge: CUP.

Johanson, L. 2003. Evidentiality in Turkic. In *Studies in Evidentiality* [Typological Studies in Language 54], A. Y. Aikhenvald & R. M. W. Dixon (eds), 273–290. Amsterdam: John Benjamins.

Klewitz, G. & Couper-Kuhlen, E. 1999. Quote-unquote? The role of prosody in the contextualization of reported speech sequences. *Interaction and Linguistic Structures*, 12: 1–34.

Köder, F. 2013. How children acquire reported speech in German and Dutch: A corpus study. In *„Perspektiven" Diskussionsforum Linguistik in Bayern/Bavarian Working Papers in Linguistics 2*, B. Sonnenhauser, C. Trautmann & P. Noel (eds), 15–28. Munich: LMU.

Lakoff, G. 1970. Linguistics and natural logic. *Synthese* 22(2): 151–271.

Lohmann, H. & Tomasello, M. 2003. The role of language in the development of false belief understanding: A training study. *Child Development* 74(4): 1130–1144.

MacWhinney, B. 2000. *The CHILDES Project. Tools for Analyzing Talk*. Hillsdale NJ: Lawrence Erlbaum Associates.

Marbach, E. S. & Yawkey, T. D. 1980. The effect of imaginative play actions on language development in five-year-old children. *Psychology in the Schools* 17: 257–263.

Nelson, K. & Gruendel, J. 1986. Children's scripts. In *Event Knowledge. Structure and Function in Development*, K. Nelson (ed.), 21–46. Hillsdale NJ: Lawrence Erlbaum Associates.

Nielsen, M. & Dissanayake, C. 2000. An investigation of pretend play, mental state terms and false belief understanding: In search of a metarepresentational link. *British Journal of Developmental Psychology* 18(4): 609–624.

Noice, H. & Noice, T. 2001. Learning dialogue with and without movement. *Memory & Cognition* 29(6): 820–827.

Nordqvist, Å. 2001a. *Speech about Speech. A Developmental Study on Form and Function of Direct and Indirect Speech*. Göteborg: Kompendiet.

Nordqvist, Å. 2001b. The use of direct and indirect speech by 1½- to 4-year-olds. *Psychology of Language and Communication* 5(1): 57–66.

Pellegrini, A. D. & Galda, L. 1982. The effects of thematic-fantasy play training on the development of children's story comprehension. *American Educational Research Journal* 19(3): 443–452.

Percus, O. & Sauerland, U. 2003. Pronoun movement in dream reports. In *Proceedings of NELS 33*, M. Kadowaki & S. Kawahara (eds), 265–284. Amherst MA: University of Massachusetts.

Plank, F. 1986. Über den Personenwechsel und den anderer deiktischer Kategorien in der wiedergegebenen Rede. *Zeitschrift für Germanistische Linguistik* 14(3): 294–308.

Recanati, F. 2000. *Oratio obliqua, oratio recta. An essay on metarepresentation*. Cambridge MA: The MIT Press.

Sanders, J. & Redeker, G. 1996. Perspective and the representation of speech and thought in narrative discourse. In *Spaces, Worlds, and Grammar*, G. Fauconnier & E. Sweetser (eds), 291–317. Chicago IL: University of Chicago Press.

Sawyer, R. K. 1996. Role voicing, gender, and age in preschool play discourse. *Discourse Processes* 22(3): 289–307.

Schank, R. C. & Abelson, R. 1977. *Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures*. Hillsdale NJ: Lawrence Erlbaum Associates.

Scott, C. L., Harris, R. J. & Rothe, A. R. 2001. Embodied cognition through improvisation improves memory for a dramatic monologue. *Discourse Processes* 31(3): 293–305.

Tannen, D. 2007. Talking Voices. *Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: CUP.

Taylor, M. 1999. *Imaginary Companions and the Children who Create Them*. Oxford: OUP.

Taylor, M. & Carlson, S. M. 1997. The relation between individual differences in fantasy and theory of mind. *Child Development* 68(3): 436–455.

Wade, E. & Clark, H. H. 1993. Reproduction and demonstration in quotations. *Journal of Memory and Language* 32(6): 805–819.

Wierzbicka, A. 1974. The semantics of direct and indirect discourse. *Papers in Linguistics* 7: 267–307.

Wilson, D. 2000. Metarepresentation in linguistic communication. In *Metarepresentations. A Multidisciplinary Perspective*, D. Sperber (ed.), 411–448. Oxford: OUP.

Wolf, D. P., Rygh, J. & Altshuler, J. 1984. Agency and experience: Actions and states in play narratives. In *Symbolic Play. The Development of Social Understanding*, I. Bretherton (ed.), 195–217. Orlando FL: Academic Press.

Yao, B., Belin, P. & Scheepers, C. 2011. Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience* 23(10): 3146–3152.

Yao, B., Belin, P. & Scheepers, C. 2012. Brain "talks over" boring quotes: Top-down activation of voice-selective areas while listening to monotonous direct speech quotations. *NeuroImage* 60(3): 1832–1842.

Yao, B. & Scheepers, C. 2011. Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition* 121(3): 447–453.

Ways of expressing action in multimodal narrations – the semiotic complexity of character viewpoint depictions

Jana Bressem[i], Silva H. Ladewig[ii] and Cornelia Müller[ii]
[i]Chemnitz University of Technology / [ii]European University Viadrina

Based on an analysis accounting for the whole body as a possible articulator in the depiction of actions, this chapter argues for an expansion of the notion of 'character viewpoint gestures' to a notion of 'multimodal action depiction from a character viewpoint'. Our study shows that speakers may deploy only single articulators, providing a semantically reduced depiction of the action, or they may deploy more bodily articulators and give a semantically rich picture of the event narrated. Our findings suggest a continuum of semiotic complexity, capturing the range of bodily involvement from less pantomimic (single articulators involved) to pantomimic depictions (more articulators involved) of actions. The paper closes by discussing our observations with respect to the notion of 'constructed action' and 'role shift' in sign languages and by giving some general remarks on the multimodal analysis of narrations.

1. Introduction

"Narrating is a communicative and reconstructive practice" (Gülich & Hausendorf 2000: 369, our translation) that assumes particular relevance and function in face-to-face interactions. By telling a (self-)experienced story, speakers can, for instance, illustrate an event, structure and sequence past events, reveal parts of their biography, gossip about others, and affirm themselves of a common social identity (Quasthoff 2000: 1295).

Narrations follow techniques of structuring discourse (e.g., opening up and closing of a narrative activity) as well as of evaluation (e.g., the creation of tension, surprise, or involvement). For these purposes, speakers have at hand a range of different linguistic devices through which narrations are developed: structuring and connecting devices mark the global structure of narratives (Gülich 1970; Quasthoff

1979), interrogative particles and connectors participate in an interactive creation of story lines, and particular stylistic and prosodic labels are used to mark respective (partial) activities (e.g., Gumperz 1992; Selting 1995). Apart from these linguistic devices, patterns of reporting and quoting are essential in narrations. In particular newer research has revealed that questions of perspective taking and the portrayal of oneself and others are constitutive for the development of narrative stories (e.g., Günthner 1996; Keim 1993).

Most research on the nature of narrations has concentrated on the abovementioned aspects and thus focused on narrations and their verbal nature (including prosody). However, when observing narrations in face-to-face interaction closely, one notices that narrations are not only staged verbally but are accompanied by bodily and gestural depictions. "Narrative language is thus not a two-dimensional affair" expressed by speech alone but rather "has a full, rounded 3-D structure, one dimension of which is imagistic" expressed in the speaker's body (Cassell & McNeill 1991: 376). Distinguishing between a narrative, metanarrative and paranarrative level, Cassell and McNeill put forward the argument that gestures take over particular functional relevance in the narration of stories as "different kinds of gestures appear, depending on where in the narratological structure the speaker is operating" (Cassell & McNeill 1991: 386). Whereas the paranarrative level is characterized by the absence of gestures or by the use of deictic gestures indicating locations in gesture space, the metanarrative level, being used for the structuring of the narrations and the story line, shows a high occurrence of metaphoric gestures, expressing abstract events and states, and deictic gestures. By far the most important level, however, is the narrative level in which iconic gestures, depicting concrete objects and events, appear. On this level, speakers have at hand several ways of depicting events when narrating a story. Imagine a person telling a story how a person drove a car against a wall. The speaker may choose to depict the event from the perspective of the character of the story, sitting in the car and crashing into the wall. In doing so, the speaker may depict the driver by holding on to the steering wheel as if driving a car and the crash by a rapid forward movement of the whole body. Such a character viewpoint gesture (McNeill 1992, 2005) expresses the event in the first-person point of view. The hand(s) and the body represent the character's hand(s) and body and the narrator of the story is "inside the gesture space" (McNeill 1992: 119) and thus inside the story. Instead of a first person view, speakers may choose the observer's viewpoint, a third-person point of view, in which the hand(s) represent one or more of the entities in the narration (McNeill 1992). In this case, the narrator may depict the wall by holding up a lateral flat hand and the crashing of the car into the wall by rapidly moving the other hand, which is formed to a fist, against the flat hand (see Figure 1; examples taken from

Figure 1. Observer viewpoint gestures, hands representing two entities: the flat hand represents a wall, the fist represents the car crashing into the wall

Ladewig, Müller & Teßendorf 2010; Müller 2010). As a third possibility, speakers may choose to combine both viewpoints and present a scene simultaneously from the perspective of the character and the observer (Cassell & McNeill 1991; Duncan 2005; McNeill, 1992; Parrill 2009; Sweetser & Stec 2013). Such "dual viewpoint gestures" (McNeill, 1992; Parrill 2009) express "simultaneously multiple perspectives on an event or scene" and may be produced "with different articulators (e.g., hands and body, hands and legs)" (Parrill 2009: 217, 276). In both-handed gestures, for instance, one hand may embody the character's viewpoint while the other hand embodies the observer's viewpoint. Parrill (2009: 279) gives the example of "a narrator who describes a character catching a ball may use character viewpoint to show the character's hand, but observer viewpoint to show the ball".

Research investigating these different types of gestural viewpoints in detail has furthermore shown that the selection of a particular viewpoint is dependent on the syntactic structure of the utterance. Character viewpoints, for instance, occur more frequently with transitive verbs and single clause sentences whereas observer viewpoints occur more frequently with intransitive or stative verbs and multi-clause sentences (Beattie & Shovelton 2001, 2002; McNeill 1992; Parrill 2010). Apart from the syntactic structure, the event structure and its spatial, imagistic, and motion properties appear to restrict the use of a particular viewpoint.

Character viewpoints occur more frequently with central events and the display of affect, whereas observer viewpoints occur more frequently with peripheral events and the display of trajectories (Kita & Özyürek 2003; McNeill 1992). Moreover, also the use of particular modalities such as the hands or the upper body appears to be restricted by viewpoint. "They may be mimetic, as when the speaker embodies the character's viewpoint. Or the speaker's hands may trace paths or assume the forms of entities being spoken about, manifesting an observer's viewpoint of their locations, actions, and movements" (Duncan 2005: 289). Accordingly, whereas character viewpoints result in mimetic representations involving more body parts, observer viewpoints seem to be restricted to the hands (Brown 2008; Casey & Emmorey 2009; Duncan 2005; Quinto-Pozos & Parrill 2012). Syntactic, lexical as well thematic differences between character and observer viewpoint gestures are assumed to be grounded in a basic difference of the two regarding the bodily anchoring and thus the conceptualization of events: Whereas character viewpoint gestures result in a depiction of events as experienced, observer viewpoint gestures depict events as observed.

Whereas existing studies mainly concentrated on a comparison of character and observer viewpoint gestures, asking for possible similarities or differences between the two, in particular with respect to their relation with speech, we suggest a different perspective in this paper. In the following sections, we will solely focus on how the body is used in narrations when depicting an action from the point of view of a character in a story. The paper addresses the following issues in particular: Based on the observation that character viewpoint gestures involve more body parts in the mimetic representation of actions, we pursue the matter of whether these depictions may show different ranges of semiotic complexity. More precisely, we attempt at shedding light on the semiotic structure of depicting actions by tracing which bodily articulators are used and how they contribute to depicting actions from a character viewpoint. In doing so, we will expand the notion of character viewpoint gestures to the idea of *multimodal action depiction from a character viewpoint*,[1] including not only the hands but also the whole body as possible semiotic resources.

With these foci, we would like to present some further insights into the complexity of multimodal depictions of actions. Although we do assume that narrations are multimodally-achieved practices in which speakers verbalize and visualize aspects of narrations (story line, actants, situations, etc.) through the

---

1. We use the term "multimodal" in this paper to refer to the involvement of different bodily resources (hand(s), head, eyes, face, and torso) in character viewpoint depictions.

interplay of speech and body, we will not investigate the interaction between the verbal and the gestural modality in creating a story but rather concentrate on the involvement of different articulators in the depiction of actions. This is in line with findings from gesture research indicating that the depiction of actions is fairly complex (e.g., Ladewig, Müller & Teßendorf 2010; Mittelberg & Waugh 2014; Streeck 2008; Teßendorf 2016). In concentrating solely on the depiction of actions, this particular phenomenon and the complexity of the bodily involvement will be analyzed in more detail.

The paper is structured as follows: First, we will discuss how we analyzed the narrations investigated in the study. Then we will introduce three examples, focusing on the deployment of bodily articulators in the depiction of actions. Based on this analysis, a continuum of semiotic complexity capturing the involvement of the different bodily articulators in narrations will be presented. The paper concludes with a discussion of some consequences for research on character viewpoint gestures and examines the findings in relation to the notion of 'constructed action' and 'role shift' in sign languages (Casey & Emmorey 2009; Cormier, Quinto-Pozos, Sevcikova & Schembri 2012; Hermann & Pendzich 2014; Herrmann & Steinbach 2012; Quinto-Pozos 2014; Quinto-Pozos & Parrill 2012 inter alia), and ends with some general remarks on a multimodal analysis of narrations.

2.   Analyzing multimodal narratives

Our study addresses the type and number of bodily articulators involved in multimodal narratives, when actions are described from a character viewpoint perspective. Articulators include, the hand(s), the head, the eyes, the face, and the torso.

Our database consists of 10 hours of audiovisual material of non-elicited interactions taken from family and friends gatherings (13 speakers participated). In this corpus, we found 31 narrations in which speakers describe actions from the viewpoint of the character performing the action. In seven out of 31 narrations only one articulator was deployed. Two articulators were involved in five multimodal depictions, three articulators were observed in five narrations. Since the use of more than three articulators results in a fairly complex miming of actions, we subsumed these instances under the category of "more than three articulators", resulting in 14 sequences of such a use of multiple articulators (see Table 1).

Table 1. Number of articulators contributing to character
viewpoint descriptions across narrations

	One articulator	Two articulators	Three articulators	More than three articulators
Narrations (31)	7	5	5	14

These character viewpoint descriptions of actions were then analyzed with respect
to which articulators participated in the depiction of the actions and to how a
variable involvement of articulators contributes to different forms and degrees of
semiotic complexity of the bodily action description. Our analysis follows three
successive steps:

Step 1   Description of articulators involved in the character viewpoint depiction
Step 2   Meaning analysis of body movement
      Step 2.1   Analysis of the meaning conveyed by each articulator involved in the
character viewpoint depiction ('meaning of the bodily form only')
      Step 2.2   Analysis of the meaning in the verbal and interactive context
('meaning of the bodily form in context')
Step 3   Determination of the degree of semiotic complexity

Our form-based linguistic approach to the study of body-movements pays particu-
lar attention to the potential differences in the articulator realizations (Bressem
& Ladewig 2011; Müller 1998, 2013; Müller, Ladewig & Bressem 2013). This take
on the analysis of body movements departs from the assumption that gestures
are motivated form gestalts, that is meaningful wholes, in which however every
aspect of a gesture's form is regarded as potentially meaningful. We therefore set
an analysis and theoretical reflection of the specifics of gesture's formal deploy-
ment center stage. For the purpose of this particular study, we have expanded
the notion of gesture forms from hand gestures to include the whole body. We
maintain, however, a close consideration of the form specifics of a bodily per-
formance with its single articulators. *In doing so, we treat each of its form aspects
as a potential meaning aspect in a multimodal depiction.* This is the core idea of a
form-based approach: any body movement is 'work', is 'effort', and we assume that
when speakers move their bodies to 'tell' parts of a narration, then every aspect of
this 'effortful movement' (to allude to Kendon's (2004) notion of 'gesture as delib-
erate expressive movement', cf. Müller 2014b) is part of the story. We suggest that
a procedure considering each body movement as a potential meaning candidate in
a multimodal description makes it possible to disentangle the semiotic structure
of character viewpoint depictions.

We have decided to include the following articulators and aspects of their formal deployment in the analysis (Table 2).

In the analysis, we start with a close description of the 'effortful' body movements involved in performing character viewpoint descriptions. In this first analytical move, we determine 'the meaning of the form'. Only in a second step, we consider how this meaning of the form, for instance somebody holding something that could be a steering wheel, is specified by its verbal and interactive context. Only then do we determine if the person acts as if steering a car or a ship, or whether he or she describes how somebody else held a steering wheel and drove a car or sailed a boat. To exemplify this methodological stance, a sample from our data will be analyzed in the following section.

Table 2. Range of articulators and aspects of formal deployment

Articulators	Aspect of form
Hands	Form features: hand shape, orientation, movement and position (Bressem 2013)
	Gestural mode of representation: acting, molding, outlining, representing (Müller 1998, 2014a)
Head	Movements of the head: e.g., shifts to the side, nodding, backwards movements (Harrison 2014; McClave 2000)
Eyes	Direction of gaze: e.g., looking sideways, looking upwards, looking at the interlocutors (e.g., Eibl-Eibesfeldt 1971; Goodwin 1981; Kendon 1967; Kendon & Ferber 1973)
Face	Facial expression: e.g., surprise, smiling, frowning (e.g., Ekman, Sorenson & Friesen 1969; Ekman, Friesen & Hager 2002; Knapp, Hall & Horgan 2013)
Torso	Movements of the upper body: e.g., shifting sideward, leaning forward or backward (e.g., Birdwhistell 1970; Frey et al. 1981; Davis 1979; Lausberg & Sloetjes 2009)
	Movements of the lower torso: e.g., abducting or adducting the leg, shifts of the pelvis, stretching the feet (e.g., Birdwhistell 1970; Frey et al. 1981; Davis 1979; Lausberg & Sloetjes 2009)

3.   Exemplification of the method

In the example "May Day", a woman is talking about her experiences made on May Day in the former GdR. In the snippet we want to focus on, the woman performs the body movement portrayed in Figure 2.

Figure 2. Body movement of the speaker in example "May Day"

Step 1.   Description of articulators involved in the character
           viewpoint depiction

The first step in the analysis is to determine which bodily articulators participate
in the depiction and to describe the particular ways in which they are moved. In
this analytical step, no verbal or contextual information is consulted. So, when we
look at the character viewpoint depiction, illustrated in Figure 2, we can see that
the right hand, the head, the eyes, the face, and the upper body participate in the
speaker's body movement. These movements can then be described by looking at
the form aspects, realized by each articulator involved in the depiction.

      As summarized in Table 3, the speaker's right hand shows the configuration of
a lax flat hand, the palm is oriented laterally and rotated once in the clockwise di-
rection. Furthermore, it is positioned in the upper right gesture space (see McNeill
1992 for the partitioning of a speaker's gesture space). The speaker's head is shifted
to the right side. While doing so, she looks forwards and slightly upwards with eyes
wide opened. She smiles and leans backwards.

Table 3. Participating articulators and their forms

Articulator	Form aspect of body movement
Hand	
– configuration:	lax flat hand
– orientation:	palm lateral
– movement type:	rotation
– movement direction:	clockwise
– position:	upper right
Head	shift to the right side
Eyes	looking forward and slightly upwards eyes wide open
Face	
– eyebrows:	raised
– facial expression:	smiling
Torso	
– upper body:	leaning backwards

Step 2.   Meaning analysis of body movements

In the following steps, the meaning conveyed by the articulators is analyzed; first without its verbal and interactional context and then in relation to speech and the interaction the speaker is participating in.

Step 2.1   *Analysis of meaning conveyed by each articulator involved in the character-viewpoint depictions*

We assume that effortful body movements are meaningful, contributing semantic or pragmatic content to a multimodal depiction. They are grounded in and abstracted from everyday actions or the perception of motions and objects (Bressem & Müller 2014; Ladewig 2014c, forthcoming; Ladewig & Bressem 2013; Müller 2010, 2014b). When looking at body movements, we can see, to a certain degree, which actions are performed or what objects are depicted. Accordingly, we believe that a basic, general meaning of body movements can be determined without knowing their specific verbal and interactional contexts. This is what we describe as 'the meaning of the bodily form'. So, for instance, in the above example we can 'see' that the person is looking forwards and slightly upwards with wide-opened eyes, looking expectantly at somebody or something. But we cannot see, whether the person or the object she is looking at, is tall or placed in the distance, nor can we determine the interactive function of eyes' movement. Now, since we assume, that every effortful body movement is potentially meaningful, we pay minute attention to the particular deployment of each articulator.

From this theoretical assumption follows a methodological procedure: each identified articulator with its instantiated form aspects is examined with regard to its meaning of the bodily form only.

Table 4. Meaning conveyed by each articulator involved
in the character-viewpoint depictions

Articulator	Form aspect	Meaning of body movement
Hand		
– configuration:	lax flat hand	Motor pattern of waving,
– orientation:	palm lateral	acting as if waving
– movement type:	rotation	
– movement direction:	clockwise	
– position:	upper right	
Head	shift to the right side	directing one's attention to the right
Eyes	looking forward and slightly upwards eyes wide open	looking expectantly at somebody or something who/which is either tall or placed in the distance
Face		
– eyebrows:	raised	pleasant anticipation of something
– facial expression:	smiling	or somebody
Torso		
– upper body:	leaning backwards	directing oneself towards something or somebody

The form features of the hand (the lax flat hand shows a lateral orientation, is rotated clockwise, and positioned in the upper right periphery) merge into a movement gestalt that can be described as an *enactment of waving*. In terms of how the hand is used to mime, the hand movement can be characterized as "acting-as-if-waving". The semiotic technique, which motivates the gesture, is "acting". In terms of Müller's notion of Gestural Modes of Representation (Müller 1998, 2014), the bodily technique of mimesis is the enactment of a body action with no instrument involved. The head's shift to the right side displays that the person in the story is oriented to the right. By looking forwards and slightly upwards with her eyes wide open, the speaker mimes looking expectantly or surprised at somebody or something, who/which is either tall or located in the distance. Together with her raised eyebrows and the smiling facial expression, pleasant anticipation of something or somebody is conveyed. By leaning backwards, the speaker shows how someone orients her/himself to an object or to a subject, which is either tall or located in the distance. By and large, the movement gestalt of waving happily at somebody or something, being either tall or set in the distance, is expressed.

Obviously such an analytical, segmental approach to the analysis of body movements can offer insights into embodied meaning construction only to a certain degree. It has to be considered as element of a movement gestalt, in which the separate contributions of the bodily articulators come together to form a holistic meaning gestalt. Notwithstanding, each bodily articulator provides its own puzzle piece to the construction of meaning and should thus be analyzed carefully for its particular contribution to the overall bodily meaning. More importantly, for our present study, this variable contribution of bodily articulators to an action description offers the ground for detecting the semiotic complexity of such bodily depictions. Speech and the broader interactional context need to be consulted in the next analytical step to arrive at a full picture of the mimed action.

Step 2.2 *Analysis of the meaning in the verbal and interactive context*
Now, we consider in which context the bodily performance that we have characterized above as "enactment of waving" is placed and how this context specifies and indexically anchors this movement gestalt.

The woman in our little example describes her experiences of a typical "May Day" in the former GdR. This was, of course, an important holiday for communist regimes and there were enormous celebrations in every city. Our speaker describes the different activities people had to attend during this day but then she shifts her focus to the more 'leisure part' of the day, namely the food court where people were free to do what they wanted.

In the example, the voice carries specific prosodic features, such as stress and loudness which evoke the meaning of speaking or calling out loud. With the sentence *und denn vorne war dis festzelt* ('and in the front there was the pavilion', see line 1 and 2 in Figure 3), the speaker introduces a particular place, which builds the locational frame for the following depiction. By breaking off the following sentence *da haste* ('there you have') she opens up a syntactic and semantic slot in which the subsequent bodily depiction is inserted. While saying *AH komm rAn hier* ('ah come here') in a loud voice she acts as if waving happily at somebody or something, being either tall or set in the distance. Taking both her utterance and her bodily depiction of action into account the speaker transmits the multimodal meaning of beckoning someone from afar. The broader interactive context informs us that the speaker re-enacts and tells how she would call her friends to eat and celebrate together at the food court, on May Day.

1. transcript	*und denn vorne war dis festzelt*	*da haste=*	**AH komm rAn hier=**
2. translation	Then in the front there was a pavillion	Then you have.	**Ah come here =**
3. deployed articulators			**right hand** **head** **eyes** **face** **upper body**

Figure 3. Transcript of the example "May Day"

Step 3. Determination of degree of semiotic complexity

Having followed the preceding analytical steps, we are now in position to precisely determine the semiotic complexity of the multimodal character-viewpoint description in our example. It involves quite a large number of bodily articulators, namely hands, head, eyes, face, and upper body, and it would count as a depiction with more than three articulators involved. In doing so, the narrator creates what can be characterized as a *pantomimic* description, i.e., a description in which a large number of bodily articulators contribute to the bodily meaning construction (Ladewig, Müller & Teßendorf 2010). As all of the deployed articulators add semantic information to the utterance meaning, a complex image of the mimed action is constructed. This is what we consider to be a high degree of semiotic complexity. The idea of degrees of semiotic complexity follows the iconic principle "more material is more meaning" (Jakobson 1966; Mayerthaler 1980; Müller & Tag

2010; Pörings & Schmitz 2003). Put differently, the picture given by this kind of pantomimic enactment is semantically richer because it is iconically richer than a simple flick of the hand can ever be.

4.   Ways of expression action in multimodal narrations

In the following section we will present examples of character viewpoint depictions of actions that illustrate the three different cases of articulator involvement introduced above: character viewpoint depictions realized bodily with one, two or three, or more articulators. We are suggesting that these depictions differ in the ways in which an event is multimodally orchestrated. Due to limitations of space, we will not spell out every single step in the analysis but focus on a description of the bodily articulators involved in expressing action within a narration. Note, however, that the steps presented in the foregoing section are necessary to identify the articulators' contributions to the overall meaning of a movement gestalt and the semiotic complexity of an action depiction.

4.1   Character viewpoint depiction involving a single articulator

The first sample shows a case of depicting actions in which only one articulator is deployed.

The sketch in Figure 4 displays a woman talking about the ritual of the family dog when visiting her grandmother. She explains what the dog used to do in order to initiate particular actions of the grandmother such as filling a bowl with water, getting food, or having the curtain moved away to look out of the window. In the part of the story telling we are focusing on now, the dog's toy box plays a major role. The speaker is describing what the dog does in order to have the dog's toy box moved out from under the closet: He runs to the closet, stands in front of it, and starts scraping.

The speaker describes the dog's action by saying *dann rennt er in flUr, kratzt.* ('Then he runs into the hallway, scrapes.' See line 1 and 2 in Figure 4.) When uttering the verb *kratzt* ("scrapes"), she uses her left flat hand with the palm oriented downwards and performs three arced movements downwards. The hand is positioned in the left periphery of the speaker's gesture space. In doing so, she looks at one of her interlocutors. The form features of the hand contribute to the movement gestalt, which can be described as an enactment of scratching. Through the verbal and interactional context we learn that the speaker describes a dog's action. It is not the speaker or another person who executes the scratching but an animal. With the

1. transcript	*Denn rennt er in flUr*	*kratzt*
2. translation	Then he runs into the hallway	**scratches**
3. deployed articulators		**left hand**
4. movement gestalt		**enact scratching**

Figure 4. Character viewpoint depiction of action involving a single articulator

verbal context in mind we can thus conclude that the hand becomes a dog's paw that executes the scratching.

Taking all possible bodily articulators into account, we see that the action is embodied by the hand only. The speaker does not use her head, face, or body to exhibit the dog's posture while scraping. The direction of gaze serves an interactive function of addressing the interlocutor. With the use of only one articulator, the bodily depiction appears iconically reduced and schematized, in a way, it can be seen as more 'abstract'. Only the speaker's left hand "stands for" the dog's action of standing in front of a closet, looking at the box, and performing a scratching movement. Thus, the action is mimed in a fairly reduced way because only one articulator is involved in miming the action and in contributing meaning to the utterance. This is what we consider an action depiction with a reduced iconicity or a low a low degree of semiotic complexity.

4.2 Character viewpoint depiction involving two articulators

In the second depiction of action (Figure 5) two bodily articulators are involved. It is taken from a conversation between several members of a family. They jointly remember their first family apartment. Having described the apartment, the mother

1. transcript	*aba denn wa der abend och lustig da ham wir denn abendbro:t*	***denn ham wa (.) was getrunken***
2. translation	But in the end the evening was funny. Then we had dinner.	**Then we had a drink.**
3. deployed articulators		**right hand head**
4. movement gestalt		**enact holding a small container and leading it towards the speaker's mouth (drinking)**

Figure 5. Character viewpoint depiction of action involving two articulators

tells a story about the visit of distant relatives, who came to examine the apartment, commenting loudly its size and furniture.

In summing up the visit, the narrator explains that the evening ended quite funny with joint drinks and a dinner. While saying *dann ham wa was getrunken* ('Then we had a drink'. See line 1 and 2 in Figure 5.), the speaker looks at her interlocutor, lifts her right hand and moves it toward her face. The form features of the hand (all fingers are bent, the palm is oriented vertically, an arced movement upwards is performed, the hand is positioned center right) merge into the movement gestalt of holding a small object, being led towards the speaker's mouth. With a short motion, the speaker's head is moved towards the speaker's hand. Taking the body movement and the verbal and interactional context together, we learn that speaker enacts drinking.

Two articulators are involved in this action depiction, namely the right hand and the head. The speaker does not use a particular facial expression or a particular posture (movement of the upper or lower body) to express the mood the characters were in. Her gaze fulfills the interactive function of addressing her interlocutor. Yet,

compared to the first example, this depiction is more complex, more iconic, and more pantomimic. Two articulators are selected by the speaker to enact the drinking action providing more mimetic 'material' and, thus, more meaning. A more complex picture of the depicted action is given, resulting in more iconic miming of the action, and displaying a higher degree of semiotic complexity.

4.3 Character viewpoint depiction involving several articulators

In the last example we see a complex multimodal depiction of an action. It is taken from a conversation between four women sharing the experiences they made when coming home from a wedding party. The story, being told by the woman illustrated in Figure 6, focuses on an incidence which took place, when the speaker, her sister, and their grandmother arrived at their apartment building. It was a funny situation, because the speaker's sister noticed that she had forgotten the key to her apartment and in order to enter the apartment nevertheless, she started climbing up the window but her grandmother would try to keep her away from doing so.

1. transcript	*oma immer*	*<<t> was*	*wILLST*	*dEnn DU*
2. translation verbatim	grandma always	what	**want**	you
2.1 translation			**(what do you want)**	
3. deployed articulators			**right hand** **left hand** **head** **eyes** **face** **upper body**	
4. movement gestalt			**enact pushing an object forward**	

1. transcript	(-)	*gEh du*	*rUnter>*
2. translation	(-)	you get	**down**
3. deployed articulators			**right hand** **left hand** **head** **eyes** **face** **upper body**
4.movement gestalt			**enact pushing an object down**

Figure 6. Character viewpoint depiction of action involving several articulators

The speaker says *und oma immer* ('and grandma always') and starts imitating her grandmother's behavior. While saying *was wILLst dEnn DU (-) gEh du rUnter* ('what do you want you get down', see line 1 to 2.1 in Figure 6) she performs two pushing movements with both hands visible in the form features flat hand, palm lateral away, straight movement away body, position center-center. The first pushing movement is performed forward, the second one is performed downwards. From the verbal and interactive context we can infer that the hands enact the grandmother's action of pushing her daughter down and hindering her to climb the wall. Additionally, she imitates her grandmother's voice (tone and stress) as well as her facial expression. She even turns her gaze away from the interlocutor to look at an imaginative subject being pushed as is done by the grandmother in the story. Here, the speaker not only uses her hands in order to construct an action but deploys several articulators, namely right and left hand, head, eyes, face, and upper body. In doing so, she provides a complex miming or pantomimic reenactment of the character's action. A large number of bodily articulators are involved, giving

a complex and iconically rich picture of the event described. We characterize this type of articulator involvement in a character viewpoint depiction as displaying a high degree of semiotic complexity.

To sum up, in all of our above examples, the speakers depict actions from the point of view of the character in the story. These actions are mimed in different ways: In the first example the speaker uses only her hand to mime the dog's scraping action. From the range of potentially available bodily resources only one articulator is selected to mime the action. As such, the action is depicted in a *less pantomimic,* less iconic way, showing a low degree of semiotic complexity. The second example is semiotically more complex and iconically richer, since two articulators are involved in the miming of drinking, namely the hand and the head. By using two articulators to contribute meaning to the utterance, a more pantomimic way of action depiction is chosen by the speaker. Here, we observe a higher degree of semiotic complexity and see an iconically richer picture. An even more complex miming of action is given in the third example in which the speaker deploys several bodily articulators simultaneously to mime the pushing action of the grandmother in the story. In this example, the speaker uses, what we characterize, a *pantomimic* enactment in which a large number of bodily articulators carry meaning, creating a more complex and iconically richer image of the action being constructed, thus showing what we conceive of as a high degree of semiotic complexity.

5.   Continuum of semiotic complexity

In the presentation of the three cases we have seen that speakers make use of different articulators when depicting actions from a character viewpoint. We have also indicated that the different articulators are involved to varying degrees in the expression of actions suggesting a continuum of iconic enrichment or of semiotic complexity (Figure 7).

Continuum of Semiotic Complexity

single articulator	*two articulators*	*more than two articulators*
less pantomimic		more pantomimic
iconically poor		iconically rich
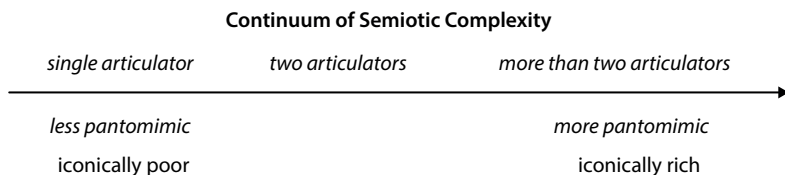
Figure 7. Continuum of semiotic complexity ranging from iconically poor to iconically rich action depictions

On the one end of the continuum, speakers may use only one articulator to depict an action. On the other end of the continuum, we find different articulators involved in the expression of actions. This range of behaviors suggests a continuum from less iconic, i.e. less pantomimic, to more iconic, i.e. more pantomimic forms of multimodal action depiction. This increasing degree of iconicity is what we conceive of as increasing semiotic complexity. It results from the higher amount of semanticized articulators, i.e. articulators that are used as 'effortful' movements in the character viewpoint depictions. These articulators are the hand(s), the head, the eyes, the face, and the torso (upper and lower body). They instantiate different form aspects, conveying a basic meaning of the bodily form that becomes specified in and through the verbal as well as interactional context. The more articulators are used simultaneously, the more form aspects are instantiated, and the more iconically rich the bodily depiction is. Moreover, a complex bodily depiction is very close to its derivational base, i.e., the original action a speaker aims to describe, because more than one articulator is usually involved in the execution of actions. The use of one articulator on the other hand is a more reduced, an iconically poorer depiction, which is more abstracted from the original action and thus more schematized.

To sum up, the continuum of semiotic complexity is based on a continuum of iconic complexity. Iconicity is here understood as the conceived similarity between a sign carrier and its meaning (Peirce 1960) and is at play in spoken and sign languages at a range of different levels and for different functions (see Jakobson 1966 and Mandel 1977, for instance). Out of a range of principles, languages particularly exploit the principle of quantity (e.g., Jakobson 1966; Mayerthaler 1980; Pörings & Schmitz 2003). Stating that something carries more form also carries more meaning, the principle is, for instance, at play, when marking the plural in spoken or sign languages. As plural forms usually consist of the stem and additional morphological material, formal complexity corresponds to conceptual complexity (more than one thing of the same kind).

"Similarity between the bodily actions we observe in others and our own perceptual and physical habits" (Mittelberg 2014: 1718) is also key to processes of sign production and reception in multimodal communication. The body in relation with speech can create different bodily icons of different semiotic substance: the entire body can function as an icon by imitating a particular posture or action or individual body parts function as an icon representing an object or action (Mittelberg 2014: 1723). Applying this to the examples discussed in this chapter, different bodily icons and moreover different degrees of iconicity are exhibited in the multimodal narrations: the more complex a bodily depiction is, that is, the more articulators are involved, the more iconic is the multimodally-expressed action. Thus, if more articulators are semanticized and take part in the action depiction, a more fine-grained

picture of the narrated action is being conveyed. Put differently and in a nutshell, the varying degrees of semiotic complexity follow the iconic principle of quantity (Bressem 2015; Müller 2008; Müller & Tag 2010).

6.   Concluding thoughts: Iconicity of action depiction and constructed action in gesture and sign language

With our discussion of multimodal depictions from naturally occurring conversations, we hope to have contributed to an expansion of the current notion and understanding of character viewpoint depictions. Our observations suggest that the ways in which character viewpoints are expressed bodily can be further differentiated according to the number of articulators involved in the narration and the respective semiotic complexity of the action depiction. As a consequence, character viewpoint depictions can be arranged on a continuum of varying semiotic complexity, ranging from less-pantomimic (performed with single articulators) to more pantomimic depictions (performed in varying combinations of hands, upper body (head, torso, facial expression, gaze and the voice)) (see also Dudis 2004, 2007 for the concept of body partitioning in sign languages). We have interpreted these observations such that the varying contributions of articulators in character viewpoint depictions go along with varying degrees of semiotic complexity: the less bodily articulators are involved, the less semiotically complex the depiction is. Following the iconic principle "more material is more meaning" (Bressem 2015; Müller 2008; Müller & Tag, 2010), a semiotically complex depiction of action is also considered iconically richer than a depiction involving only a single articulator. As a matter of fact, the image created by a complex bodily depiction is much closer to the gestural base, that is, the original action, as more significant parts of the action's structure are preserved. In cases in which only a small number of articulators participate in action depictions, the speaker picks out only a few salient parts to stand for the whole action. The image created is, thus, more schematized, more abstract and looks, in a way, more 'linguistic'. Since sign languages make use of the same 'articulatory material' as gestures, it is interesting to compare gesturally mimed actions to linguistic iconicity in sign languages. Our observations on the continuum of semiotic complexity suggest that gesture and sign resemble each other in the different aspects of iconic mappings, which Taub has described in her analogue-building model for sign languages: i.e. "image selection", "schematization", and "encoding" (Taub 2001).

> Both imitator and language user start with a concept to communicate, select a sensory image to represent that concept, and create an analogue of that image in some modality (visual, auditory, temporal, etc.); in both cases [e.g., linguistic iconicity and miming; remark by the authors], the end product preserves significant parts of the structure of the original image. (Taub 2001: 57)

However, Taub points out, that in the case of the gestural miming of actions, the "element of *conventionalization* is lacking", at least in spontaneous gestural creations (ibid., italics in the original). This means that a linguistic system putting constraints on the gestural sign created by the speaker is missing. On the contrary, in sign language the image to be created has to meet the language-specific system constraints and, thus, has to undergo a "linguistic schematization process" (ibid.).

But, when we consider, the construction of action in sign languages, things look differently, meaning that in the reporting of another's actions – a phenomenon known as 'action role shift' or 'constructed action' (Casey & Emmorey 2009; Cormier, Quinto-Pozos, Sevcikova & Schembri 2012; Hermann & Pendzich 2014; Herrmann & Steinbach 2012; Quinto-Pozos 2014; Quinto-Pozos & Parrill 2012 inter alia) – expressions can range from rather gestural in nature to more abstract and thus more linguistic (Quinto-Pozos 2007). This leads to the question of the linguistic status of constructed action. Does the reproduction of another person's action count as gesture or as sign? The answer given by Wilcox and Xavier (2013) suggests that "constructed action usage events[2] […] may begin life as gestural depictions". Their "repeated use by signers in certain contexts and genres leads to schematization" (Wilcox & Xavier 2013: 100). Moreover, symbolically less complex depictions in sign language, which are comparable to what we have characterised as semiotically-poorer, more abstract, more schematized depictions, tend to become a lexical item by undergoing schematization processes. On the other hand "[s]ymbolically-complex constructed actions resist lexicalization because the variation across usage events is too great to develop a low-level, lexical schema." (ibid.: 100)

From a point of view of linguistic gesture analysis, we would like to suggest that similar schematization or conventionalization processes can be observed in gestures, meaning that semiotically-less complex, i.e. more abstract, more schematized depictions of actions can become entrenched, forming "structural islands" of conventionalized gestures (Müller, Bressem & Ladewig 2013). An example would, for instance, be the family of 'away gestures', which are all based on manual actions that share the effect of a cleared body space (throwing away, sweeping away,

---

2.   Wilcox & Xavier (2013) follow a cognitive-linguistic take on the investigation of signs and gesture and conceive tokens of gesture and sign as "usage events" (Langacker 2008).

brushing away, holding away) and which share in the semantic themes of negative assessment, refusal, and negation (Bressem & Müller 2014; on gestures of negation see also Harrison 2009 and Kendon 2004). As this example shows, these "recurrent gestures" (Ladewig 2010, 2014a; Müller 2010), as we call them, may fulfill other duties in discourse than the "mere" iconic depiction of a character's actions (ibid., Bressem & Müller 2014; Kendon 1992, 2004; Müller 2004; Teßendorf 2014, 2016 inter alia). They often serve discourse-related functions such as negative assessment, refusal and negation, or they may function as topic or the comment markers of an utterance (see e.g., Kendon 2004; Seyfeddinipur 2004).

The parallels between gesture and sign are certainly due to the fact that gesture and sign appear to be "fashioned from the same material", to quote Adam Kendon (2000: 49) or that both make use of the same expressive medium (Ladewig & Bressem 2013; Müller 2013, Müller, Bressem & Ladewig 2013). These structural similarities furthermore indicate that character viewpoint depictions may function as a source for signs (e.g., Cormier et al. 2012; von Loon, Pfau & Steinbach 2014; Wilcox, Rossini & Pizuto Antinoro 2010). In any case, these observations certainly underline that a discussion of a possible gestural or bodily basis of constructed actions or role shift in sign languages is a worthwhile enterprise indeed.

What are the implications of our observations for researching multimodal narrations? First of all, the systematic partitioning of the body in different modalities and articulators as documented in this paper provides the grounds for a detailed reconstruction of how speakers use speech and body when narrating a story. By broadening the scope from language only to the use of multimodal articulation of narration a different perspective is offered (see Parrill 2009). Our observations underline that in contrast to verbal narrations, which due to the character of the medium "speech" are bound to linearity and to a successive representation of actions, the simultaneous use of multiple articulators allows for different and more complex ways of narrating a story. How complex the interrelation of verbal and visual modalities in narrations told by hearing speakers is, is particularly obvious in the examples in which multiple strategies are used simultaneously. We believe therefor that the phenomena presented in this chapter are not only of interest for gesture research or a possible comparison of gestures with sign languages but also for the investigation of narrations in general, in which questions of perspective taking and the portrayal of oneself and others play, in fact, central roles (e.g., Günthner 1996; Keim 1993). The disentangling of the multimodal expressions of actions in narrations offers the possibility of gaining further knowledge on the multimodal and embodied meaning creation in narrations.

Acknowledgements

References

Beattie, G. & Shovelton, H. 2001. How gesture viewpoint influences what information decoders receive from iconic gestures. In *Oralite et gestualite: Interactions et comportements multimodaux dans la communication*, C. Cave, I. Guaitella & S. Santi (eds), 283–287. Paris: L'Harmattan.

Beattie, G. & Shovelton, H. 2002. An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology* 93(2): 179–192.

Birdwhistell, R. 1970. *Kinesics and Context. Essays on Body, Motion, Communication*. Philadelphia PA: University of Pennsylvania Press.

Bressem, J. 2013. A linguistic perspective on the notation of form features in gestures. In Müller et al. (eds), 1079–1098.

Bressem, J. 2015. Repetition als Mittel der Musterbildung bei redebegleitenden Gesten. In *Handbuch Satz, Äußerung, Schema*, C. Dürscheid & J. G. Schneider (eds), 422–442. Berlin: Walter de Gruyter.

Bressem, J., Ladewig, S. H. & Müller, C. 2013. *Linguistic Annotation System for Gestures (LASG)*. In Müller et al. (eds), 1098–1125.

Bressem, J. & Ladewig, S. H. 2011. Rethinking gesture phases – Articulatory features of gestural movement? *Semiotica* 184(1–4): 53–91.

Bressem, J. & Müller, C. 2014. *The family of AWAY-gestures: Negation, refusal and negative assessement*. In Müller et al. (eds), 1592–1605.

Brown, A. 2008. Gesture viewpoint in Japanese and English: Cross-linguistic interactions between two languages in one speaker. *Gesture* 8(2): 256–276.

Bühler, K. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart: Fischer.

Casey, S. & Emmorey, K. 2009. Co-speech gesture in bimodal bilinguals. *Language and Cognitive Processes* 24(2): 290–312.

Cassell, J. & McNeill, D. 1991. Gesture and the poetics of prose. *Poetics Today* 12(3): 375–404.

Cormier, K., Quinto-Pozos, D., Sevcikova, Z. & Schembri, A. 2012. Lexicalisation and delexicalisation processes in sign languages: Comparing depicting constructions and viewpoint gestures. *Language and Communication* 32: 329–348.

Davis, M. 1979. Laban analysis of nonverbal communication. In *Nonverbal Communication: Readings with Commentary*, S. Weitz (ed.), 182–206. Oxford: OUP.

Dudis, P. G. 2004. Body partitioning and real-space blends. *Cognitive Linguistics* 15(2): 223–238.

Dudis, P. G. 2007. *Types of depiction in ASL*. Ms, Gallaudet University. <http://www.gallaudet.edu/faculty-staff/linguistics/dudis_paul.html>

Duncan, S. 2005. Gesture in signing: A case study in Taiwan Sign Language. *Language and Linguistics* 6(2): 279–318.

Eibl-Eibesfeldt, I. 1971. Transcultural patterns of ritualized contact behavior. In *Behavior and Environment. The Use of Space by Animals and Men*, A. H. Esser (ed.), 238–246. New York NY: Plenum.

Ekman, P., Sorenson, E. R. & Friesen, W. 1969. Pan-cultural elements in facial displays of emotions. *Science* 164(3875): 86–88.

Ekman, P., Friesen, W. & Hager, J. C. 2002. *Facial Action Coding System*. Salt Lake City UT: Research Nexus.

Frey, S., Hirsbrunner, H. P., Pool, J. & Daw, W. 1981. Das Berner System zur Untersuchung nonverbaler Interaktion: I. Die Erhebung des Rohdatenprotokolls; II. Die Auswertung von Zeitreihen visuell-auditiver Information. In *Methoden der Analyse von Face-to-Face-Situationen*, P. Winkler (ed.), 203–268. Stuttgart: Metzler.

Goodwin, C. 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press New York.

Gumperz, J. 1992. Contextualization revisited. In *The Contextualization of Language* [Pragmatics & Beyond New Series 22], P. Auer & A. di Luzio (eds), 39–54. Amsterdam: John Benjamins.

Gülich, E. 1970. *Makrosyntax der Gliederungssignale im gesprochenen Französisch*. Munich: Fink.

Gülich, E. & Hausendorf, H. 2000. Vertextungsmuster Narration. In *Text-und Gesprächslinguistik/Linguistics of Text and Conversation*, K. Brinker, G. Antos, W. Heinemann & S. F. Sager (eds), 369–385. Berlin: Walter de Gruyter.

Günthner, S. 1996. The prosodic contextualization of moral work: An analysis of reproaches in 'why'-formats. In *Prosody in Conversation*, E. Couper-Kuhlen & M. Selting (eds), 271–302. Cambridge: CUP.

Harrison, S. 2009. The expression of negation through grammar and gesture. In *Studies in Language and Cognition*, J. Zlatev, M. Andren, M. J. Falck & C. Lundmark (eds), 405–409. Cambridge: CUP.

Harrison, S. 2014. *Head shakes: Variation in form, function, and cultural distribution of a head movement related to "no"*. In Müller et al. (eds), 1496–1501.

Herrmann, A. & Pendzich, N.-K. 2014. *Nonmanual gestures in sign languages*. In Müller et al. (eds), 2149–2163.

Herrmann, A. & Steinbach, M. 2012. Quotation in sign languages. In *Quotatives: Cross-linguistic and Cross-disciplinary Perspectives* [Converging Evidence in Language and Communication Research 15], I. van Alphen & I. Buchstaller (eds), 203–228. Amsterdam: John Benjamins.

Jakobson, R. 1966. *Selected writings*. Berlin: Mouton de Gruyter.

Keim, I. 1993. Sprachvariation als konstitutives Merkmal eines sozialen Stils. Am Beispiel einer innerstädtischen Welt „Kleiner Leute". *Soziolinguistik. Der Deutschunterricht* 3: 44–56.

Kendon, A. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26: 22–63.

Kendon, A. 1980. Gesture and speech: Two aspects of the process of utterance. In *Nonverbal Communication and Language*, M. R. Key (ed.), 207–227. The Hague: Mouton.

Kendon, A. 1992. Some recent work from Italy on quotable gestures (emblems). *Journal of Linguistic Anthropology* 2(1): 92–108.

Kendon, A. 2000. Language and gesture: Unity or duality. In *Language and Gesture*, D. McNeill (ed.), 47–63. Cambridge: CUP.

Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge: CUP.

Kendon, A. & Ferber, A. 1973. A description of some human greetings. In *Comparative Ecology and Behavior of Primates*, R. P. Michael & J. H. Crook (eds), 591–668. New York NY: Academic Press.

Kita, S. & Özyürek, A. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48: 16–32.

Knapp, M., Hall, J. & Horgan, T. 2013. *Nonverbal Communication in Human Interaction*. Boston MA: Wadsworth Cengage Learning.

Ladewig, S. H. 2010. Beschreiben, suchen und auffordern – Varianten einer rekurrenten Geste. *Sprache und Literatur* 41(105): 89–111.

Ladewig, S. H. 2011. Putting the cyclic gesture on a cognitive basis. *CogniTextes* 6. <http://cognitextes.revues.org/406> (June 2013).

Ladewig, S. H. 2014a. *Recurrent gestures*. In Müller et al. (eds), 1558–1575.

Ladewig, S. H. 2014b. *The cyclic gesture*. In Müller et al. (eds), 1605–1618.

Ladewig, S. H. 2014c. *Creating multimodal utterance. The linear integration of gestures into speech*. In: Müller et al. (eds), 1662–1677.

Ladewig, S. H. (forthcoming). *Integrating Gestures – cognitive grammar multimodal*. Berlin: De Gruyter Mouton.

Ladewig, S. H. & Bressem, J. 2013. New insights into the medium hand – Discovering structures in gestures based on the four parameters of sign language. *Semiotica* 197: 203–231.

Langacker, R. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: OUP.

Lausberg, H. & Sloetjes, H. 2009. Coding gestural behavior with the NEUROGES – ELAN system. *Behavioral Research Methods* 41(3): 841–849.

van Loon, E., Pfau, R. & Steinbach, M. 2014. *The grammaticalization of gestures in sign languages*. In Müller et al. (eds), 2133–2149.

Mandel, M. 1977. Iconic devices in American Sign Language. In *On the other hand. New perspectives on American Sign Language*, L. A. Friedman (ed.), 57–108. New York: Academic Press.

Mayerthaler, W. 1980. Ikonismus in der Morphologie. *Zeitschrift für Semiotik* 2: 19–37.

McClave, E. Z. 2000. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics* 32: 855–878.

McNeill, D. 1992. *Hand and Mind. What Gestures Reveal about Thought*. Chicago IL: University of Chicago Press.

McNeill, D. 2005. *Gesture and Thought*. Chicago IL: University of Chicago Press.

Mittelberg, I. 2006. *Metaphor and Metonymy in Language and Gesture: Discursive Evidence for Multimodal Models of Grammar*. PhD dissertation, Cornell University. Ann Arbor MI: ProQuest/UMI. <https://search.proquest.com/docview/305321684>

Mittelberg, I. 2014. *Gestures and iconicity*. In Müller et al. (eds), 1712–1732.

Mittelberg, I. & Waugh, L. 2014. *Gestures and metonymy*. In Müller et al. (eds), 1747–1766.

Müller, C. 1998. *Redebegleitende Gesten: Kulturgeschichte, Theorie, Sprachvergleich*. Berlin: Arno Spitz.

Müller, C. 2004. Forms and uses of the Palm Up Open Hand. A case of a gesture family? In *Semantics and Pragmatics of Everyday Gestures*, C. Müller & R. Posner (eds), 233–256. Berlin: Weidler Verlag.

Müller, C. 2008. *Metaphors. Dead and Alive, Sleeping and Waking. A Dynamic View*. Chicago, University of Chicago Press.

Müller, C. 2010. Wie Gesten bedeuten. Eine kognitiv-linguistische und sequenzanalytische Perspektive. *Sprache und Literatur* 41(105): 37–68.

Müller, C. 2013. *Gestures as a medium of expression: The linguistic potential of gestures*. In Müller et al. (eds), 202–217.

Müller, C. 2014a. *Gestural modes of representation as techniques of depiction*. In Müller et al. (eds), 1677–1687. Berlin: De Gruyter Mouton.

Müller, C. 2014b. Gesture as "deliberate expressive movement." In *From Gesture in Conversation to Visible Action as Utterance: Essays in Honor of Adam Kendon*, M. Seyfeddinipur & M. Gullberg (eds), 127–152. Amsterdam: John Benjamins.

Müller, C., Bressem, J. & Ladewig, S. H. 2013. *Towards a grammar of gesture: A form-based view*. In Müller et al. (eds), 707–733.

Müller, C., Ladewig, S. H. & Bressem, J. 2013. *Gestures and speech from a linguistic perspective: A new field and its history*. In Müller et al. (eds), 55–81.

Müller, C., Ladewig, S. H. & Teßendorf, S. 2010. Structures of meaning and reference – Investigating gestures of the concrete and abstract. *Paper presented at the 4th Conference of the International Society for Gesture Studies*, Frankfurt/Oder, Germany.

Müller, C. & Tag, S. 2010. The dynamics of metaphor: Foregrounding and activating metaphoricity in conversational interaction. *Cognitive Semiotics* 6: 85–120.

Müller, C., Cienki, A., Fricke, E., Ladewig, S. H., McNeill, D. & Teßendorf, S. (eds). 2013. *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Volume 1. Berlin, Boston: De Gruyter Mouton.

Müller, C., Cienki, A., Fricke, E., Ladewig, S. H., McNeill, D. & Bressem, J. (eds). 2014. *Body-Language-Communication: An International Handbook on Multimodality in Human Interaction*. Volume 2. Berlin, Boston: De Gruyter: Mouton.

Parrill, F. 2009. Dual viewpoint gestures. *Gesture* 9(3): 271–289.

Parrill, F. 2010. Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5): 650–668.

Peirce, C. S. 1960. *Collected Papers of Charles Sanders Peirce (1931–1958)*, Vol. I: *Principles of Philosophy,* Vol. II: *Elements of Logic*, C. Hartshorne & P. Weiss (eds). Cambridge MA: The Belknap of Harvard University Press.

Pfau, R. & Steinbach, M. 2006. Modality-independent and modality specific aspects of grammaticalization in sign languages. *Linguistics in Potsdam* 24: 3–98. <www.ling.uni-potsdam.de/lip>

Pörings, R. & Schmitz, U. 2003. *Sprache und Sprachwissenschaft. Eine kognitiv orientierte Einführung*. Tübingen: Gunter Narr.

Quasthoff, U. 1979. Konversationelle Erzählungen im Sozialamt: Ihr Einsatz als kommunikative Strategien. In *Soziologische Analysen. Referate aus den Veranstaltungen der Sektionen der deutschen Gesellschaft für Soziologie und der ad-hoc-Gruppen beim 19. Deutschen Soziologentag*, R. Mackensen & F. Sagebiel (eds), 515–534. Berlin: Deutsche Gesellschaft für Soziologie (DGS).

Quasthoff, U. 2000. Erzählen als interaktive Gesprächsstruktur. In *Text- und Gesprächslinguistik/ Linguistics of Text and Conversation*, K. Brinker, G. Antos, W. Heinemann & S. F. Sager (eds), 1293–1309. Berlin: Walter de Gruyter.

Quinto-Pozos, D. 2007. Why does constructed action seem obligatory? An analysis of "classifiers" and the lack of articulator-referent correspondence. *Sign Language Studies* 7(4): 458.

Quinto-Pozos, D. 2014. Enactment as a (signed) language communicative strategy. In Müller et al. (eds), 2163–2170.

Quinto-Pozos, D. & Parrill, F. 2012. Comparing viewpoint strategies used by co-speech gesturers and signers. *Paper presented at the 5th Conference of the International Society for Gesture Studies*, Lund, Sweden.

Selting, M. 1995. Sprechstile als Kontextualisierungshinweise. In *Stilfragen. Jahrbuch 1994 des Instituts für deutsche Sprache*, G. Stickel (ed.), 225–256. Berlin: Walter de Gruyter.

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, E., Gilles, P., Günther, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peter, L., Quasthoff, U. M., Schütte, W., Stukenbrock A. & Uhmann, S. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online Zeitschrift zur verbalen Interaktion* 10: 353–402.

Seyfeddinipur, M. 2004. Meta-discursive gestures from Iran: Some uses of the 'Pistol Hand'. In *The Semantics and Pragmatics of Everyday Gestures*, C. Müller & R. Posner (eds), 205–216. Berlin: Weidler.

Streeck, J. 2008. Depicting by gesture. *Gesture* 8(3): 285–301.

Sweetser, E. 2012. Introduction: Viewpoint and perspective in language and gesture, from the Ground down. In *Viewpoint in Language: A Multimodal Perspective*, B. Dancygier & E. Sweetser (eds), 1–22. Cambridge: CUP.

Sweetser, E. & Stec, K. 2013. Managing multiple viewpoints: Coordinating embedded perspective in multimodal narratives. *Paper presented at the ICLC-12*, Edmonton, Canada.

Taub, S. 2001. *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge: CUP.

Teßendorf, S. 2014. *Pragmatic and metaphoric gestures – Combining functional with cognitive approaches*. In Müller et al. (eds), 1540–1558.

Teßendorf, S. 2016. Actions as sources of gestures. In *Beyond Language Boundaries: Multimodal Use in Multilingual Contexts*, M. Fernández-Villanueva & K. Jungbluth (eds), 34–54. Berlin, Boston: De Gruyter.

Wilcox, S. 2004. Gesture and language: Cross-linguistic and historical data from signed languages. *Gesture* 4(2): 43–75.

Wilcox, S. 2007. Routes from gesture to language. In *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*, E. Pizzuto, P. Pientrandrea & R. Simone (eds), 107–137. Berlin: Walter de Gruyter.

Wilcox, S., Rossini P. & Antinoro E. P. 2010. Grammaticalization in sign languages. In *Sign Languages*, D. Brentari (ed.), 332–355. Cambridge: CUP.

Wilcox, S. & Xavier A. N. 2013. A framework for unifying spoken language, signed language, and gesture. *Todas as Letras-Revista de Língua e Literatura* 15(1): 88–110.

Nominal referential values of semantic classifiers and role shift in signed narratives

Gemma Barberà and Josep Quer

CNRS/Paris 8 and ICREA/Universitat Pompeu Fabra

Bringing together the areas of dynamic semantics and signed discourse, this article focuses on the dynamic potential of referring expressions, such as semantic and limb classifiers, and role shift constructions. On the basis of the Catalan Sign Language version of the Aesop's fables, a qualitative analysis is presented, which focuses on the interaction between referring expressions and the accessibility scale. While the incorporation of semantic and limb classifiers constructions into the accessibility hierarchy leads to a more fine-grained analysis, role shift also reveals itself as an essential mechanism to associate the classifier handshape to the corresponding discourse referent. Dynamic operations are taken into account with the main aim of offering a novel proposal on discourse accessibility structure in signed narrative discourse.

1. Introduction

As natural languages, sign languages are not instantiated through isolated sentences but rather through signed discourses, as larger relevant units of communication. Discourse is considered to be an entity that is constantly being built and updated. Irrespective of the language modality, some referring expressions have the potential of changing and updating the context, whereas some others interconnect the referring expressions vis-à-vis the entities in the discourse context and, in principle, do not update it because they do not contain descriptive material. Therefore, and according to the general view, discourse is constructed by two operations (von Heusinger 2007). The first one is based on the context change potential achieved through the update of the accessibility structure of a discourse and implemented by definite and indefinite noun phrases (NPs). The second one resides in the context-dependent interpretation implemented by anaphoric pronouns. While the first operation updates and manipulates the list of possible referents, the second one implements and evidences the updates through the use of coreferential expressions.

These two dynamic operations are essential for the construction and development of any kind of discourse, regardless of the language modality. The knowledge on signed discourse currently available is still very limited and this article aims at enlarging it by offering an innovative description of semantic classifiers and an account for their referential value in signed discourse.

This article investigates the dynamic potential of different types of referring expressions in signed discourse, with a special focus on semantic and limb classifiers and role shift constructions, and how they interact in the accessibility scale. The analysis is framed in terms of dynamic semantics, in particular within von Heusinger's (2007) account of discourse accessibility structure. This investigation is based on naturalistic data and it presents a qualitative analysis of Catalan Sign Language (*llengua de signes catalana*, LSC) narratives, and more concretely of a number of Aesop's fables, signed by two native deaf signers. The proposal of the present contribution is two-fold. On the one hand, we propose that accessibility in signed discourse is structured across semantic relations, where semantic and limb classifiers and role shift constructions are incorporated into the accessibility hierarchy, leading to a more fine-grained analysis. On the other hand, we propose that semantic classifiers are proforms, which involve two dynamic operations: the interpretation of the *classifier handshape* is dependent on the previously established discourse referent (they thus undertake the same function as pronouns); but at the same time, the *classifier construction* has a context change potential, since it includes descriptive content as part of a full predication. As will be shown, the set of all possible denotations for a classifier handshape is defined as a class determined by its physical and geometrical properties. However, that set is drastically reduced by its anchoring to the most salient antecedent that precedes it and by being combined with the predicational root of the classifier construction. In our fragment of LSC narratives, role shift is an essential mechanism, necessary to associate the classifier handshape to the corresponding discourse referent.

The article is organised as follows. In Section 2, the tools to understand dynamic semantics and accessibility theories are offered. Previous work on reference tracking in signed languages is reviewed and the fine-grained account of accessibility on which this article is based is presented. In Section 3, semantic classifiers found in LSC narratives are presented by focusing mainly on their coarticulation with role shift. The two discourse functions they have, namely the anaphoric and the predicative one, are explained in detail. In Section 4, the proposal is extended to role shift constructions serving as link between antecedents and anaphors instantiated by semantic and limb classifiers. In Section 5, the main findings of the article are summarised.

2.   Background

2.1   Dynamic semantics and salience

It has been generally accepted that sentences within a discourse are interpreted with respect to their truth conditions, but they also need to be interpreted in connection with the context. Every new sentence introduced into a discourse is connected to the preceding ones, but at the same time it adds information and increments the context. The context is thus constantly changing and formal and traditional theories of discourse treat sentences as denoting functions from contexts to contexts (Heim 1982; Kamp & Reyle 1993). What may be called 'dynamic' about dynamic semantics is that it involves a notion of interpretation that contributes some kind of change. In dynamic semantic theories, the meaning of an expression is said to be its 'context-change potential' (Nouwen 2003). The building blocks of discourse are referring expressions, which do not only weave it by interconnecting the various linguistic elements but they also pick up discourse referents of the universe of discourse, i.e. the objects of thought the conversation is about. The different kinds of referring terms display different properties. Full NPs and proper names have reference independently and they select a specific discourse referent from the universe of discourse. In contrast, pronouns and demonstratives do not inherently select a discourse referent from the universe of discourse. Rather, they restrict the entities to which they may refer to. An example of this is the proper name *Joana*, which rigidly picks up the discourse referent for "Joana", which is ontologically connected to the real human being known by that name. Hence, between the proper name and the discourse referent there is a direct reference connection. In contrast, pronouns have some features that allow restricting the amount of potential antecedents. The specific features of English pronouns, for instance, are gender, number and case. For example, the pronoun *she* has the features [+ feminine], [+ singular], and [+ nominative]. *She* does not directly identify the specific discourse referent to which it refers, but it rather selects a subgroup of possible entities from the discourse domain, namely those that are compatible with [+ feminine], [+ singular], and [+ nominative]. Once the selection is done, syntactic and pragmatic constraints allow picking up the corresponding entity among the set of possible discourse referents. In this article we focus on underspecified referring expressions, such as pronominal forms. They are characterised by two main features: (i) they inherently specify some properties of the discourse referent and they act as set restrictor devices among the entities from the universe of discourse, and (ii) their referential interpretation is dependent on a prominent discourse referent they pick up.

Recently, discourse studies have shown that the degree of prominence of a discourse referent directly affects the referring term that will be chosen to denote

such an entity.[1] The form chosen reflects the salience of the entity within a specific fragment and, according to the literature, pronominal and weaker forms are the expressions used when the entity is taken to be actively salient (i.e. prominent) in the consciousness of the addressee (different but related approaches have been proposed, such as the Assumed Familiarity by Prince (1981); the Givenness Hierarchy by Gundel, Hedberg & Zacharski (1993), and the Accessibility Marking Scale by Ariel (1990); for overviews of the psycholinguistic literature on discourse referent accessibility and anaphora processing, see Almor (1999) and Arnold (2010), among others.). Salience theories offer a procedural analysis of referring expressions, as marking varying degrees of mental accessibility. The basic idea is that referring expressions instruct the addressee to retrieve a certain piece of given information from his memory by indicating how accessible this piece of information is to the addressee at the current stage of discourse. This is shown in the following examples, when uttered out of the blue. In a situation where two teachers are discussing about how intelligent their students are, Example (1) can only be uttered when the student has been mentioned in preceding discourse or also when the student is present in the physical context. In both cases the discourse referent being talked about is very salient. In contrast, (2) will only be uttered in a context where the discourse referent is not prominent. Therefore, the sender uses a definite NP and a relative clause to provide more information about the referent being talked about.

(1)   She is very smart.

(2)   The student we met yesterday evening is very smart.

One important contribution of these approaches is that they take into account that natural languages provide senders with the means to *encode* the salience of the discourse referent to the addressee. The general prediction is that when discourse referents are not prominent or distant from the potentially anaphoric expressions, a relatively lower salience marker will be chosen, as shown in (2). When the referent is prominent and/or very recently mentioned in the co-text, it will be coded by a relatively high accessibility marker, as seen in (1). Therefore, the accessibility hierarchy is articulated in a scale where NPs formed by a full noun phrase with a modifier are considered to be low accessibility markers, and verbal person inflections and null arguments, depending on the language, are considered to be high accessibility

---

1.   This article uses the following terminology: *salience* is used as a generic term to refer to the fact of having different degrees of mental prominence when talking about discourse referents; *accessibility* is a linguistic concept to refer to the relation between the degree of salience of a discourse referent and the choice of referring expression; *salience spreading* is a concept taken from von Heusinger (2007) to refer to a function of the discourse that yields different ordered sets corresponding to the descriptive content of the referring expression (see Section 2.3).

markers. The present contribution proposes to incorporate classifier handshapes[2] and role shift constructions into the accessibility hierarchy, by dealing with the face-to-face interaction that characterises sign language and the corresponding oral modality, as opposed to the written modality on which accessibility studies have been previously based.

2.2   Previous work on reference tracking in sign languages

Research into the mechanisms displayed in sign languages for reference tracking has been so far rather limited, and it has partly concentrated on their acquisition. In this section a brief overview of the results of previous works will be presented as background for our study.

With a focus on the development of narrative skills in children acquiring British Sign Language (BSL), Morgan (2006) establishes a simple hierarchy of referring expressions in terms of explicitness, that is, of how much descriptive content they encode and, consequently, how transparent they are in the identification of its referent or antecedent. His study considers the role of NPs, entity classifiers and role shift in the introduction, reintroduction and maintenance of discourse referents (what Morgan calls referential function), and it follows the view that NPs are the most explicit ones and role shift the least explicit, with entity classifiers somewhere in between in terms of explicitness.

Putting aside the acquisitional pattern, if we concentrate on the results about the distribution of the three types of expressions in the adult control group, it becomes clear that the basic distinctions surface quite neatly. NPs are almost always used for introducing referents, while only half of the time for reintroducing them and almost never for maintaining reference. Role shift is mostly used to maintain reference, followed by entity classifiers, and even less used to reintroduce referents. In this data, role shift never has the introducing function, and there are only some instances of entity classifiers that serve this goal. Morgan and Woll's (2003) study of the acquisition of body classifiers, which we would roughly subsume under role shift, also show that they are mainly employed to maintain a referent in discourse.

---

2. Classifier handshapes in sign languages are bound morphemes that are affixed to abstract verb roots and are linked to verb arguments. They share many properties with verbal classifiers in spoken languages, namely (i) they are affixes attached to a verb stem, (ii) they are linked to the subject or object argument of the verb, (iii) they may have an anaphoric function, (iv) they allow variability in the choice of classifier, among others (see Zwitserlood 2012 for a deeper comparison between verbal classifiers in spoken and sign languages; see Section 3.1 in this article for a description of semantic classifiers in sign languages).

One of the few studies addressing the distribution of referring expressions in sign language discourse is Swabey (2002). She relies on the Givenness Hierarchy proposed by Gundel et al. (1993), which is an implicational scale of cognitive statuses that are marked by different forms of referring expressions. Those cognitive statuses refer to memory and attention state of the addressee and range from most to least restrictive as far as the intended sets of entities are concerned. Thus, for instance, referents 'in focus' are at the current center of attention and are conveyed by zero or unstressed pronominals. At the opposite end of the scale, we find referents that are 'type identifiable', that is recoverable from the type conveyed by the description, but the addressee does not need to have any specific referent in mind. In English, NPs with the indefinite article *a* typically encode this type of cognitive status.

Table 1. The givenness hierarchy (Gundel et al. 1993)

In focus	Activated	Familiar	Uniquely identifiable	Referential	Type identifiable
it	*that, this, this* N	*that* N	*the* N	indefinite *this* N	*a* N

In her analysis of ASL narrative re-tellings, Swabey found that there were no instances of classifiers used to refer to referents with the status at most 'type identifiable' or at most 'uniquely identifiable'. For referents with the status at most 'referential', classifiers were always accompanied by a preceding noun to specify the referent of the classifier. In these cases the noun and corresponding classifier were counted as one referring expression. The majority of the classifiers used for at most 'familiar' and at most 'activated' referents usually had accompanying specifying nouns. The exception to this was classifiers that were used repeatedly and always to track the same referent. For referents that were in focus the noun for the classifier did not have to be specified. On the other hand, nouns were used for referents that had a familiar or a referential status. These results are in accordance with the shared view that the more salient a referent is, the less descriptive content is used to refer to it. Kibrik and Prozorova's (2007) work on reference tracking in Russian Sign Language within Accessibility Theory reaches very similar conclusions, but they do not include classifiers or role shift among the reference tracking mechanisms they examine.

In sign language research, the intuition that classifiers function as proforms, that is, as markers that stand for the noun and have some referential properties while combined with a predicate has been present from the early sign language research (Engberg-Pedersen & Pedersen 1985; Friedman 1975; Garcia & Sallandre 2013; Herrero 2004; Kegl 1986; Kegl & Wilbur 1976). However, no account has

been proposed from this perspective. Some studies have simply assumed that an anaphoric relation between the classifier handshape and an argument of the predication is present (Benedicto & Brentari 2004; Chang, Su & Tai 2005; Cuxac 2000; Glück & Pfau 1998; Zwitserlood 2003, 2012). In these accounts, movement or localisation in the construction is taken to be a verb or root stem. The classifier as well as the locus in space are considered functional elements, such as inflectional affixes.

2.3   A more complex account of salience

The form of a referring expression is standardly taken to be a reflection of the degree of salience of the discourse referent it is linked to. Thus, while an indefinite NP introduces a discourse-new referent, a pronoun anaphorically takes us back to an already introduced one or to a non-introduced, but very salient referent. Definite NPs are also used to refer to introduced referents, but unlike pronouns, they are chosen when the referent is less salient, which explains their richer descriptive content. This general view has been embraced as the basic principle of discourse coherence and it has been developed mostly in different, but related accessibility approaches, as previously presented in Section 2.1.

However, according to von Heusinger (2007), "accessibility" of a discourse referent is not the only determining factor for the choice of referential expression. The relation between the antecedent and the discourse anaphor is determined by the accessibility of the discourse, i.e. the access of the anaphoric term to the discourse item, on the one hand, and the ranking of different discourse items, on the other. Different aspects contribute to accessibility:

i.   activation or accessibility status (depending on lexical type, descriptive content and syntactic function of the anaphor);
ii.   accessibility relation (distance and syntactic structure between the anaphoric expression and its antecedent, unity, context knowledge, encyclopaedic knowledge, inferential knowledge);
iii. accessibility hierarchy (informativity of the referring expression);
iv.   accessibility structure (syntactic structure, discourse structure, whereby we can assign an ordered set of accessible entities to a discourse domain or segment)
v.   salience of the objects in some model (salience is a property of a set associated with descriptive material expressed in a referring expression).

For von Heusinger, the accessibility structure of a discourse is formed by listing the salient items of each set associated with some predicate used in that discourse, and the relations between the listed sets (hyponymy, hyperonymy …), thus relying not

only on the degree of context dependency of a referring expression, but also on its potential for changing the context.

Discourse theories like Discourse Representation Theory (DRT, Kamp & Reyle 1993) and Centering Theory (CT, Grosz, Joshi & Weinstein 1995) approach accessibility in different ways, but they are not able to cover all the required mechanisms. In DRT accessibility is not gradable: an anaphoric relation is represented as an identification of the new discourse referent with an accessible one. The discourse referents form a set of accessible antecedents with respect to a discourse domain. Discourse referents are simply listed. In CT, the attentional state represents the availability of discourse referents at any given point in the discourse. It assumes fine-tuning among accessible discourse items, which is mirrored in the ranking of accessible items. It provides strategies to find antecedents for anaphoric pronouns in discourse segments, but cannot account for the antecedent of an NP, which depends on global accessibility.

Von Heusinger (2007) exploits the role of the descriptive content of the expression that introduces the discourse item in order to overcome the pitfalls in other approaches mentioned above. Most theories focus on the use of pronouns, which do not have significant descriptive material, and ignore nominal descriptions, with considerable descriptive content. Definite NPs are typically interpreted as static terms in the general view. However, von Heusinger explores the role of anaphoric expressions with descriptive content like definites, which change the accessibility structure. He establishes that accessibility is not the function of a discourse that yields a single set of accessible elements, but rather a function of the discourse that yields different ordered sets corresponding to the descriptive content of the referring expression. Thus, an expression like "the small bird" in English does not only make the set [Small_Birds: {a}] salient, but also the corresponding supersets [Birds: {a}] and [animate objects: {a}].[3] An expression not only changes the most-accessible element of the set introduced, but also that of some of the relevant supersets of this set, through the process labelled *salience spreading*. By virtue of this, a subsequent expression "the bird" can pick up the same referent by linking to the superset [Birds: {a}]. A pronoun "he" can also pick up on the same referent by being linked to the biggest superset [animate objects: {a}].

From this perspective, accessible discourse items are sets ranked with respect to the predicate by which they are introduced or activated, rather than a single set of ranked elements without further association to the predicate. As already mentioned at the beginning, this article investigates the dynamic potential of different types of referential expressions in signed discourse, with a focus on semantic classifiers and role shift, and how they interact in the process of salience spreading.

---

3.   This notation is taken over from von Heusinger's (2007) paper.

3. Referring expressions in signed narratives

3.1   Semantic classifiers

Signed narratives show a rich array of referring expressions, which contribute to building and developing discourse. Besides nouns and index signs directed to signing space, Catalan Sign Language (LSC) makes a great use of semantic classifiers and role shift constructions to refer to the discourse referents involved. Under the label *semantic classifiers*, both entity and limb classifiers are considered to be included here. Entity classifiers represent a broad class of noun objects (Supalla 1986). The classification of handshapes is established according to visual and geometrical properties of the entity. The predicate indicates movement or location of the entity denoted. Limb classifiers represent body parts of the entity denoted (Supalla 1986). The predicate is realised as the movement root. Figure 1 shows an instance of an entity classifier. The signer uses the 3-handshape, which denotes a legged animal with more than two legs. The classifier construction, formed by the nominal and the movement predicate, denotes a lying down event.

Figure 1. Entity classifier handshape

In Figure 2 we observe two different handshapes functioning as entity classifiers. One is expressed through the dominant hand and it is articulated with the thumb-handshape and denotes a long upright entity. The other, expressed through the non-dominant hand, is articulated with a flat-handshape and denotes a flat entity, without much volume. The three entity classifier handshapes shown in Figure 1 and Figure 2 represent a broad class of possible noun denotations. They do not directly refer to a particular discourse referent, but rather inherit the semantic properties of the previously introduced antecedent. In this case, the three-handshape (Figure 1) and the thumb-handshape (Figure 2) stand for the hare. The flat-handshape (Figure 2) stands for the turtle.

Figure 3 shows an example of a limb classifier. The flat-handshape stands for the feet of the turtle. The movement of the classifier is the predicational root and

Figure 2. Entity classifier handshapes for two discourse referents

denotes the movement of the feet. Note that the handshape for the entity classifier in Figure 2 and the handshape for the limb classifier in Figure 3 is exactly the same, represented with a flat-handshape. Importantly, the predication (movement) in both classifier constructions disambiguates the kind of classifier and the meaning attributed. Moreover, the limb classifier is coarticulated with the role shift of the entity denoted, as indicated by the facial expression of the signer. As shown in Section 4, the combination of limb classifier and role shift restrict the domain of interpretation of the entity referred to.
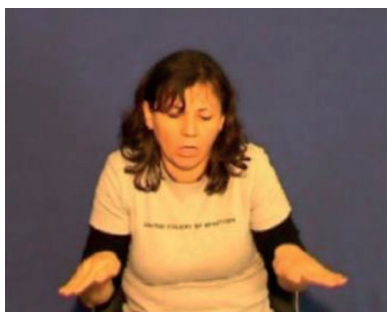
Figure 3. Limb classifier

Interestingly, classifier handshapes include three main nominal features: (i) referentiality, (ii) anaphoric potential, and (iii) definiteness. As for referentiality, we have just seen in the previous examples that classifier handshapes establish a link to a discourse referent from the domain of interpretation. This dependent referentiality is shown in the inherent anaphoric potential: they are dependent on the antecedent previously introduced, which provides the link with the referent from the domain. Last but not least, they are definite in terms of the familiarity conditions they present. Since they are attached to an antecedent present in the universe of discourse they cannot refer to a new, unknown entity. Classifier handshapes are embedded

in classifier constructions, where not only the entity but also the predicate are expressed. These constructions may also sometimes be embedded, at the same time, under role shift constructions. This is precisely the topic of Section 3.3.

3.2 Double function of classifiers

Within signed discourse, both entity classifiers and limb classifiers show a double function. First, when considering classifier handshapes, they have an anaphoric function. That is, the handshape is coreferential with a previously introduced discourse referent, as seen in Section 3.1. But when considering the classifier construction as a whole, a pure predicative function arises. Let us first focus on the anaphoric function. Semantic classifiers are contextually dependent on the previously introduced discourse referent. As underspecified forms, they need an antecedent introduced earlier in order to receive an interpretation. The corresponding simplified Discourse Representation Structure of (3) is shown in (4). The nouns "rabbit" and "turtle" introduce each a discourse referent and this is represented with a corresponding variable appearing in the discourse domain, that is the upper box. They are respectively represented as x and y. The predicates appearing in the fragment in (3) are represented in (4) as conditions updating the context. A Discourse Representation Structure is true if there is a function that maps the free variables (the upper box) to entities in the world such that the conditions (the lower box) are satisfied (see Nouwen 2003 for a deeper exposition). The classifier handshapes also introduce new variables (z, u, and v), but they need to be equated to a previous discourse referent. In this fragment, for instance, the discourse referent for "rabbit" is referred to through the bare noun, the thumb-handshape entity classifier and the Q-handshape limb classifier.[4]

(3) RABBIT TURTLE CLe(thumb):entity-standing/CLe(B):entity-standing

$$\overline{\qquad\qquad\text{rs:rabbit}\qquad\qquad}\qquad\qquad\overline{\qquad\qquad\text{rs:rabbit}\qquad\qquad}$$

CLe(thumb):run-forward /CLe(B):move-slow SPEED CLl(Q):legs-moving

'There was a rabbit and a turtle. The rabbit started to run forward fast, very fast, and moving its legs, while the turtle advanced very slowly.'

---

4. We follow the usual glossing conventions in the sign language literature, according to which manual signs are represented by the capitalised word corresponding to the translation of the sign. The relevant abbreviations for the purposes of this paper are the following: IX# (index pointing sign; the numbers refer to the grammatical person); #-VERB-# (verb agreeing with subject and object); subindices mark direction towards sign space: l (low), u (up), ip (ipsilateral); cl (contralateral); ce (centre); a (spatial location establishment) and also binding relations (i); CL for classifier constructions indicating classifier type (entity, limb …), handshape in parentheses and rough meaning description. A line above the glosses indicates the scope of nonmanuals: br (brow raise); sq (squinted eyes), rs (role shift). Reduplication of signs is indicated by +++.

(4)

x, y, z, u, v
rabbit (x)
turtle (y)
CLe(thumb):run-forward (z)
z = x
CLe(B):move-slow (u)
u = y
CLl(Q):legs-moving (v)
v = x

The three variables (i.e. the logic constructs that are identified with discourse referents) point to the particular entities already introduced in this fragment of discourse. This is why in the simplified Discourse Representation Structure in (4), the variables are identified under the identity equation, which is obtained by means of suitability motivations based on morphosyntactic and also pragmatic criteria. From a morphosyntactic point of view, the equation between the variable introduced by semantic classifiers and the variable introduced by the corresponding noun (i.e. the antecedent) is obtained through matching of geometrical features. Previous research has shown that classifier handshapes may be analysed as agreement markers (Supalla 1986; Glück & Pfau 1998; Zwitserlood 2003; Zwitserlood & van Gijn 2006). However, from a semantic-pragmatic point of view, role shift plays a crucial role in the equation between the variables introduced by this kind of referring expressions, as will be discussed in Section 4.1.

Interestingly, the antecedent is not always introduced previous to the anaphoric expression. In some contexts, it is also possible that an underspecified classifier handshape appears without previous introduction of the coreferential noun. That is, instances of backwards anaphora (i.e. cataphora) are also felicitous with entity classifiers. However, it is important to note that these cases are mainly restricted to literary contexts and only possible with classifiers that happen to be quite lexicalised. In the fragment in (5), the underspecified entity classifier denoting a two-legged entity is uttered first. In the subsequent sentence, the coreferential chain is established and the discourse referent attributing meaning to the anaphora is uttered.

(5) PARC IX SEAT BENCH CLe(N):flat-surface/CLe(2):legged-entity-seated.
         _br_
     MAN NEWSPAPER READ.
     'In a park, there was someone seated on a bench. It was a man reading the newspaper.'

Semantic classifiers, just like pronouns and unlike nouns, do not have rich descriptive content. They do not change the accessibility structure associated to nominal referents. They only do it as part of a full predication; that is, if the whole classifier

construction is considered. The second function we propose here is based on the notion of *Context Change Potential*, first presented by Heim (1982). According to Heim, the meaning of a sentence is no longer a set of static truth-conditions. Rather, the meaning of a sentence is best viewed in terms of its context change potential. A sentence meaning is a function from contexts to contexts; that is, it can be uttered in a certain class of contexts, and produces certain possible changes in those contexts as a result. Technically, Heim (1982: 294) defines the context change potential as a function that assigns to every context F the resulting context F′, which is brought about by uttering condition Φ in a situation in which F obtains. In the formula below, F′ is the value of the context change potential of Φ for the argument F.

(6) $F + \Phi = F'$

Thus, the second function that classifiers have is that of making the context updates evident. However, a context update is only possible if we consider the classifier construction as a whole and include the handshape and the predicate movement in the operation. As shown below, this complex construction introduces a new condition to the common ground, which has the potential of updating the previous context.

   To understand this more clearly, let us pick up our previous Example (3) and focus on the semantic representation already shown in (4), here repeated as (7) for convenience. As already indicated, the variable that introduced each discourse referent in the discourse domain is instantiated by uttering a noun and also by uttering a semantic classifier. As a matter of fact, the handshape classifier is part of a full predication, which in its turn adds a new condition in the discourse domain. In (7) these conditions are instantiated by the predicates "run-forward", "move-slow" and "legs-moving". In subsequent sentences, new conditions are added and contribute to incrementing the context.

(7)

x, y, z, u, v
rabbit (x)
turtle (y)
CLe(thumb):run-forward (z)
z = x
CLe(B):move-slow (u)
u = y
CLl(Q):legs-moving (v)
v = x

As a matter of interest, semantic classifiers not only update the context but they also have the potential of restricting anaphoric relations. Classifiers are inserted in main clauses, but in some contexts they may be also inserted in relative clauses, as full predications. In the following example, we observe a classifier construction

that functions as a main clause predication whose subject features a relative clause containing another classifier construction. The non-manuals are a crucial marking of relative clauses. In (8) the relative clause is coarticulated with brow raise and squinted eyes denoting a familiar entity already mentioned in the discourse. The classifier construction inserted in the relative clause contributes to the context update, but at the same time restricts the number of possible potential antecedents already present in the common ground.

(8) $\overline{\text{[MAN IX3 TREE HIDE CLe(2):seated-entity]}}^{\text{br, sq}}_{\text{rel}}$ CLe(2):climb-down

     'The man who was hiding on top of the tree climbed down.'

3.3   Coarticulation of role shift and semantic classifiers

The so-called role shift construction is the genuine means of sign languages used to convey the utterances or thoughts ascribed to a discourse agent. Previous studies have interestingly shown the coarticulation of role shift and classifiers constructions, arguing for a complex interaction in signed utterances (Earis & Cormier 2013; Garcia & Sallandre 2013; Perniss 2007; Quinto-Pozos 2007). The following fragment shows an instance of coarticulation of role shift construction with limb handshape classifiers. The discourse referent is introduced with a noun and it is then referred back to by the body-parts expressed with limb classifiers (Q-handshape) coarticulated with role shift, as shown by the horizontal line in the glosses, which represents the scope of the coarticulation.

(9)   RABBIT YES AGREE

     $\overline{\text{CLl(Q):legs-moving SPEED CLl(Q):legs-moving CLgen:mov}}^{\text{rs:rabbit}}$

     $\overline{\text{EFFORT MAXIMUM CLgen:mov CLl(Q):legs-moving}}^{\text{rs:rabbit}}$

     'The rabbit agreed and he started running extremely fast.'
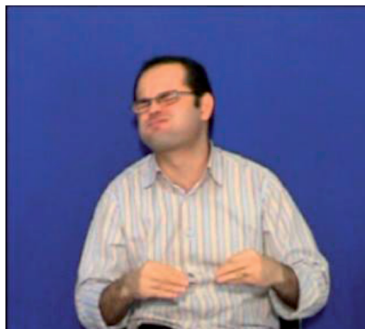
Figure 4. Limb classifier coarticulated with role shift

In our data, role shift is mainly combined with limb classifiers. However, instances of role shift co-occurring with entity classifiers are also possible (Perniss 2007). The following excerpt is an instance of the discourse referent for "rabbit" referred back to by two handshape classifiers denoting the whole entity: the thumb- (Figure 5) and the 3-handshape (Figure 6). As shown in the horizontal line in the glosses, the articulation of the two classifier handshapes co-occurs with particular movements by the body of the signer, identified as role shift. In these fragments, the body of the signer reproduces the movements of the rabbit, simultaneously depicted as a whole entity at the manual level.

Figure 5. Entity classifier (thumb-handshape) coarticulated with role shift

Figure 6. Entity classifier (3-handshape) coarticulated with role shift

(10) COINCIDE DAY HEAT STRONG SUN RABBIT FEEL START TIRED
                                                      rs:rabbit
      CLe(thumb):entity-moving SEARCH SEE TREE SMALL
                                          rs:rabbit
      CLe(3):entity-moving CLe(3):lie-down.
      'It was a very hot and sunny day. The rabbit was running and he lay down next to a small tree.'

4. Accessibility in signed anaphoric chains

In the discussion so far it has been pointed out that the set of all possible denotations for a classifier handshape is defined as a class determined by its geometrical properties. However, that set is drastically reduced by its anchoring to the most salient antecedent that precedes it and by being combined with the predicational root of the classifier construction. In this section we will concentrate on the role of yet another type of referring expression, namely limb classifiers, in establishing anaphoric chains in signed discourse and how they combine with the other type of expressions examined so far in von Heusinger's approach to salience spreading, previously presented in Section 2.3.

4.1 Licensing the identity equation and associative anaphora

An important property for the interpretation of classifiers in anaphoric chains is the function that role shift plays in the connection between the antecedent and the anaphor. When presenting the anaphoric function of semantic classifiers (Section 3.2), we have seen that the link between the classifier and the corresponding antecedent is obtained by means of suitability motivations based on morphosyntactic features. However, as already pointed out, from a semantic-pragmatic point of view, role shift plays a crucial role in resolving the equation between the variables. In fact, role shift licenses the identity equation of the referent linked to the semantic classifier (the variable it contributes) with a previously introduced discourse referent. The data shows that role shift is necessary to associate the semantic classifier handshape with the corresponding discourse referent. Indeed, from a discourse perspective geometrical feature matching is not enough for the interpretation of classifiers. Limb classifiers functioning as such help us illustrate this, for unlike entity classifiers, they are always coarticulated with role shift: otherwise they are interpreted as entity classifiers referring to a limb that is not just part of the animate referent (e.g. turtle vs. a turtle's leg). Even the same handshape is ambiguous with respect to the kind of discourse referent it stands for. For instance, in our LSC fable examples the same Q-handshape is associated to a hare (legs) and to a bear (snout) and can only be disambiguated by the role shift coarticulated with it, which contributes to making the anaphoric chain explicit. Similarly, as already showed in Section 3.1, the same flat-handshape may be used as an entity classifier denoting a turtle (Figure 2, repeated here for convenience as Figure 7) and a limb classifier denoting the legs of the turtle (Figure 3, repeated here as Figure 8). As a matter of fact, while only the second case appears compulsorily coarticulated with role shift, this is not obligatory in the entity classifier case.

Figure 7. Flat-handshape instantiating an entity classifier
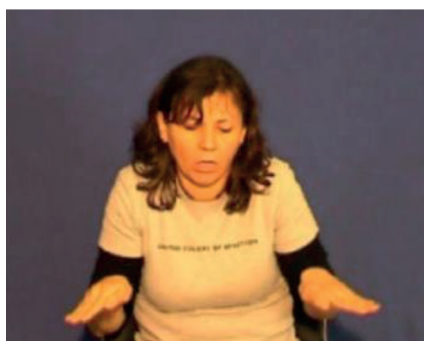
Figure 8. Flat-handshape instantiating a limb classifier

The mechanism at play with limb classifiers as anaphoric elements is a whole-part relationship, which requires an accommodation operation in order to understand the dependency between the anaphor and the antecedent. It is an instance of bridging or associative anaphora based on encyclopaedic knowledge, that is one of the cases that von Heusinger (2007) mentions as instances of indirect anaphora. Indirect anaphora crucially rely on additional information that is present in the accessibility structure.

When addressees interpret a definite associative NP, they must take into account that all definite NPs carry a presupposition of existential uniqueness. This implies that addressees can access an entity that is presented as the only one of the type expressed by the noun of the definite NP. In the case of associative NPs that introduce a new entity (NP2), this entity must be easily accessible through, for instance, a definite NP1. Take for instance the sequence "I just listened to a symphony this morning, but the composer escapes me". It is general knowledge that symphonies are composed by a musician and, even if the identity of the author is

unknown to the speaker, a referent for it is introduced by a definite description by means of associative anaphora.

Stereotypical part-whole relations are a case in point where indirect coreferentiality exists between the entities designated by NP1 and NP2. In such contexts, the existential uniqueness of the entity referred to by the NP2 is motivated by the introduction of another entity in the prior discourse. In this case, the definite NP2 is felt to be an incomplete description whose referential interpretation calls anaphorically for the linguistic context. Here, referential interpretation implies that the addressee establishes what referent in the model it should be related to. For such uses, in English examples, the definite determiner entails that the entity referred to by the NP serving as a trigger is not only connected coreferentially, but also salient (Charolles 1999). In the case of LSC, the NP serving as an antecedent to the associative limb classifier does not provide direct access to the referent; it only provides a pointer leading to it following an inferential and accommodation procedure.[5] However, the link is overtly instantiated by the use of role shift scoping with the articulation of the classifier construction.

4.2 Accessibility and semantic relations

As shown in the salience spreading account (Section 2.3), previous accessibility studies have been proven to focus only on the use of pronouns, which do not have significant descriptive material, and ignore nominal descriptions, with considerable descriptive content. Definite descriptions do not only contribute to the context with descriptive content but also with the semantic relations established among them, which change the accessibility structure. Let us make a first attempt at understanding how each element of the anaphoric chain contributes to it on the basis of a concrete example like (11). This fragment reproduces the passage of the fable "The bear and the two travellers", where the two friends see the bear approaching them. While the signer is still in the role of the most active discourse referents (the two friends), a new one is lexically introduced with the bare noun BEAR and followed by an entity classifier. This classifier is anchored to the superset of animates that the noun BEAR has triggered through salience spreading. Next, role shift switches to the bear-referent and it gets coarticulated with a limb classifier that represents the bear arms. As a whole, these two expressions are jointly anchored to the same set of animate entities made salient by the first mention of the bear-referent through a noun, as represented in Figure 9.

---

5. As pointed out by an anonymous reviewer, the bridging features of the two kinds of semantic classifiers (entity and limb) resemble the weak/strong distinction of definite articles attested cross-linguistically (Schwarz 2013). However, determining whether a uniqueness or a familiarity analysis applies to each kind of classifier is beyond the scope of this article.

$$\overline{\text{br}} \quad \overline{\text{rs:bear}} \quad \overline{\text{rs:bear}}$$

SEE BEAR$_1$ CLe(1):entity-approaching$_2$ CL1(Q):bear's-arms$_3$

{bears: {a}}

{animals: {a}}

{animates: {a}}

{entities: {a}} 1=2=3
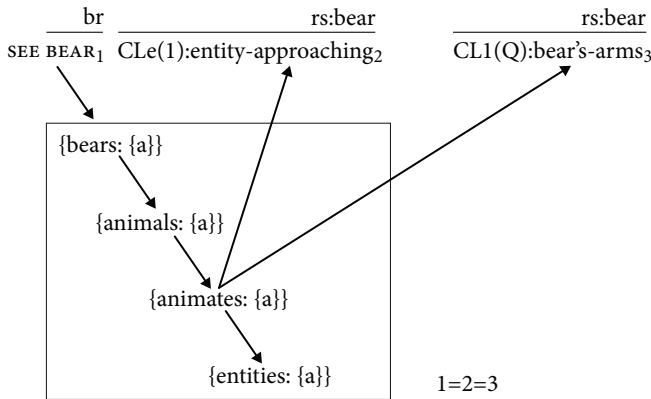
Figure 9. Salience spreading and anaphoric chain of Example (11)

(11) $\overline{\text{br}}$ $\overline{\text{rs:bear}}$ $\overline{\text{rs:bear}}$
SEE BEAR$_1$ CLe(1):entity-approaching$_2$ CLl(Q):bear's-arms$_3$
'They saw that the bear was approaching, walking in their direction.'

Following Kegl (1986), we include role shift marking as a central part of the anaphoric chain, materializing subject clitics. The way she characterises it is as follows:

> Role prominence seems closely tied with first-person point of view and is generally restricted to NPs which are [+ human]. Non-human or inanimate role-prominent NPs are possible, but they become personified. (…) The RP-CL [role prominence clitic] serves a dual function. Besides marking role prominence, it functions as a subject clitic. Role prominence is associated with subject positions only. It is also invariably singular. (Kegl 1986: 289)

Rather than [+ human], we would argue that role shift is tied to [+ animate] arguments: role shift will always place the discourse referent at the level of "animate objects" in the salience spreading proposal in von Heusinger (2007: 141), through a process of personification of animals and objects, as we see in the above examples.

In combination with a classifier, role shift is sufficient to change the salience of discourse referents. This property becomes most evident in cases of role shift in constructed dialogue, where the alternating interlocutors are simply signalled by role shift. In this respect, the dynamic contribution of role shift extends beyond that of adding or activating a discourse referent: it changes the accessibility structure at a given point in discourse.[6] Thus, in Example (12), the second friend in the story is reintroduced after a sequence where the bear was the most salient discourse referent. This is achieved by articulating an entity classifier with role shift.

---

6. We thank an anonymous reviewer for highlighting the importance of this property of role shift in the context of our discussion.

(12) pro_{bear} $\overline{\text{CLl(Q):snout-sniffing "IX3 DEAD ALREADY IX3, BAH!"}}^{\text{rs:bear}}$CLl:bear's arms

LEAVE FOREST CLe(1):direction LEAVE $\overline{\text{CLe(N):lie-down}}^{\text{rs:friend2}}$

'The bear sniffed the man and thought "He's already dead, bah!". He walked away and left entering the forest. The man was lying down.'

Given these assumptions and Kegl's analysis, we take role shift as part and parcel of the anaphoric chain in the salience-spreading representation. Its contribution is mediated by an empty subject pronoun that is coreferential with the intended discourse referent. In Figure 10 we adapt the previous representation correspondingly for Example (11).

$\overline{\text{SEE BEAR}_1}^{\text{br}}$ $\overline{\text{pro}_2\text{CLe(1): entity-approaching}_3}^{\text{rs:bear(2)}}$ $\overline{\text{pro}_4\text{CL1(Q): bear's-arms}_5}^{\text{rs:bear(4)}}$

{bears: {a}}

{animals: {a}}

{animates: {a}}

{entities: {a}}   1=2=3=4=5
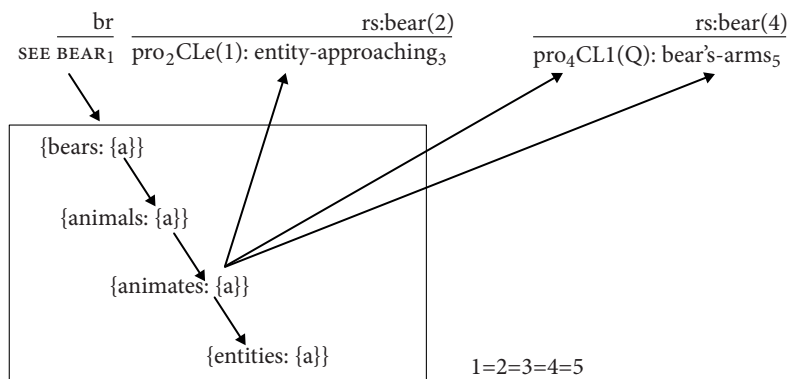
Figure 10. Modified salience spreading and anaphoric chain of Example (11)

Although limb classifiers and role shift tend to be coreferential, this is not always the case, for they can be anchored to different discourse referents, as in the example represented in Figure 11, where the role corresponds to the lying friend and the limb classifier stands for the bear's snout while sniffing him.

Figure 11. Bear sniffing lying friend

This raises the issue of having simultaneous activation of discourse referents and it triggers the question whether a discourse referent is more prominent than the other and, if so, which one. In this specific example, the man is activated by role shift, while the bear is activated by the limb classifier. We would like to argue that role is the most prominent argument by definition and, as just argued above, it turns its associated referent into the most salient one. A concrete possibility for the example at hand in Figure 11 would be to interpret it as a simultaneous passive-like construction, as suggested in one way or another in the few works that address the existence of passive constructions in sign languages (Wilbur 1987; Kegl 1990; Saeed & Leeson 1999; Janzen et al. 2001; Hansen 2007). However, resolving this question is beyond the scope of this study and we leave it for future research.

5.   Conclusions

This article has presented a novel account of the interaction between semantic and limb classifiers and role shift in LSC narratives. It has proven that these referring expressions play a crucial role in establishing anaphoric chains in signed discourse, much like pronouns. The set of entities from the discourse domain to which semantic classifiers are linked is importantly reduced by its anchoring to the most salient antecedent that precedes it and by being combined with the predicational root of the classifier construction. Therefore, besides their anaphoric function, semantic classifiers as part of classifier constructions also have the potential of updating and incrementing the context and thus contributing to its dynamic nature.

We have also seen that a static view of accessibility is insufficient to account for the distribution of referring terms in general and that a more complex account including semantic relations among referring expressions provides a fairer analysis of the dynamic coherence in natural discourse production. More concretely, salience spreading accounts naturally for the distribution of intertwined referring expressions and contributes to the accessibility structure update of narrative signed discourse. From an analytical point of view, limb classifiers have been presented here as associative anaphora connected to a previously introduced discourse referent. The link between these highly underspecified anaphoric elements and the corresponding antecedent is licensed by role shift constructions coarticulated with semantic classifiers, and especially with limb classifiers. The data shows that role shift is necessary to associate the semantic classifier handshape with the corresponding discourse referent. Role shift has thus been argued to license the identity equation of the referent linked to the semantic or limb classifier with a previously introduced discourse referent, and to turn it into the most prominent at a given point in discourse.

The proposal presented here is an innovative contribution to the young field of the semantics-pragmatics interface in sign language and more concretely to the study of referring expressions in narrative signed discourse. This study opens up new issues to delve into the semantics-pragmatics interface, such as the simultaneous activation of different discourse referents and the corresponding salience ranking and definiteness constraints. Moreover, other interesting issues may be also tackled from a prosodic and from a syntactic perspective: do the handshape classifiers belong to the same clause as the antecedent or are they part of different clauses? If they are taken to belong to a different clause as suggested by the non-manuals coarticulated, an analysis of handshape classifiers functioning as resumptive pronouns for a left-dislocated noun phrase (Glück & Pfau 1998) could be pursued. These syntactic-prosodic issues open up a promising avenue of research that should be pursued in future work.

Acknowledgements

References

Almor, A. 1999. Constraints and mechanisms in theories of anaphor processing. In *Architectures and Mechanisms for Language Processing*, M. Crocker, M. Pickering & C. Clifton (eds), 341–354. Cambridge: CUP.

Ariel, M. 1990. *Accessing Noun-phrase Antecedents*. London: Routledge.

Arnold, J. E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4(4): 187–203.

Benedicto, E. & Brentari, D. 2004. Where did all the arguments go? Argument-changing properties of classifiers in ASL. *Natural Language and Linguistic Theory* 22(4): 743–810.

Chang, J., Su, S. & Tai, J. H.-Y. 2005. Classifier predicates reanalysed, with special reference to Taiwan Sign Language. *Language and Linguistics* 6(2): 247–278.

Charolles, M. 1999. Associative anaphora and its interpretation. *Journal of Pragmatics* 31: 311–326.

Cuxac, C. 2000. *La Langue des Signes Française (LSF): Les voies de l'iconicité*. Paris: Ophrys.

Earis, H. & Cormier, K. 2013. Point of view in British Sign Language and spoken English narrative discourse: The example of 'The Tortoise and the Hare'. *Language and Cognition* 5(4): 314–343.

Engberg-Pedersen, E. & Pedersen, A. 1985. Proforms in Danish Sign Language. Their use in figurative meaning. In *SLR'83 Proceedings of the III International Symposium on Sign Language Research*, W. Stokoe & V. Volterra (eds), 202–210. Silver Spring MD: Linstok Press.

Friedman, L. 1975. Space, time, and person reference in American Sign Language. *Language* 51(4): 940–961.

Garcia, B. & Sallandre M.-A. 2013. Reference resolution in French Sign Language (LSF). In *Crosslinguistic Studies on Noun Phrase Structure and Reference* [Syntax and Semantics 39], P. Cabredo Hofherr & A. Zribi-Hertz (eds), 316–364. Leiden: Brill.

Glück, S. & Pfau, R. 1998. On classifying classification as a class of inflection in German Sign Language. In *ConSole VI Proceedings*, T. Cambier-Langeveld, A. Lipták & M. Redford (eds), 59–74. Leiden: Sole.

Grosz, B., Joshi, A. & Weinstein, S. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 2(21): 203–225.

Gundel, J., Hedberg, N. & Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274–307.

Hansen, M. 2007. Why can German Sign Language (DGS) do without a passive construction? Ways of marking semantic roles in DGS. Dissertation abstract. *Sign Language & Linguistics* 10(2): 213–222.

Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD dissertation, University of Massachusetts, Amherst.

Herrero, Á. 2004. Una aproximación morfológica a las construcciones clasificatorias en la lengua de signos española. *ELUA Estudios de lingüística* 18: 151–186.

von Heusinger, K. 2007. Accessibility and definite noun phrases. In *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference* [Studies in Language Companion Series 86], M. Schwarz-Friesel, M. Consten & M. Knees (eds), 123–144. Amsterdam: John Benjamins.

Janzen, T., O'Dea, B. & Shaffer, B. 2001. The construal of events: Passives in American Sign Language. *Sign Language Studies* 1(3): 281–310.

Kamp, H. & Reyle, U. 1993. *From Discourse to Logic. Introduction to Model Theoretic semantics of Natural Language*, Formal Logic and Discourse Representation Theory. Dordrecht: Kluwer.

Kegl, J. 1986. Clitics in American Sign Language. In *The Syntax of Pronominal Clitics* [Syntax and Semantics 19], H. Borer (ed.), 285–309. New York, NY: Academic Press.

Kegl, J. 1990. Predicate argument structure and verb-class organization in the American Sign Language lexicon. In *Sign Language Research. Theoretical Issues*, C. Lucas (ed.), 149–176. Washington DC: Gallaudet University Press.

Kegl, J. & Wilbur, R. 1976. When does structure stop and style begin? Syntax, morphology and phonology vs. stylistic variation in American Sign Language. *Papers from the Annual Regional Meeting, Chicago Linguistic Society* 12: 376–396.

Kibrik, A. A. & Prozorova, E. V. 2007. Referential choice in signed and spoken languages. In *Proceedings of 6th DAARC*, A. Branco, T. McEnery, R. Mitkov & F. Silva (eds), 41–46. Porto: Centro de Linguistica da Universidade do Porto.

Morgan, G. 2006. The development of narrative skills in British Sign Language. In *Advances in Sign Language Development in Deaf Children*, B. Schick, M. Marschark & P. Spencer (eds), 314–343. Oxford: OUP.

Morgan, G. & Woll, B. 2003. The development of reference switching enconded through body classifiers in British Sign Language. In *Perspectives on Classifier Constructions in Sign Languages*, K. Emmorey (ed.), 297–310. Mahwah NJ: Lawrence Erlbaum Associates.

Nouwen, R. 2003. *Plural Pronominal Anaphora in Context: Dynamic Aspects of Quantification*. PhD dissertation, Utrecht University.

Perniss, P. 2007. *Space and Iconicity in German Sign Language (DGS)*. PhD dissertation, Radboud University.

Prince, E. 1981 Toward a taxonomy of given-new information. In *Radical Pragmatics*, P. Cole (ed.), 223–256. New York NY: Academic Press.

Quinto-Pozos, D. 2007. Why does constructed action seem obligatory? An analysis of classifiers and the lack of articulator-referent correspondence. *Sign Language Studies* 7(4): 458–506.

Saeed, J. & Leeson, L. 1999. Detransitivisation in Irish Sign Language. *Paper presented at the European Science Foundation Intersign Meeting on Morphosyntax*. Siena, Italy, March.

Schwarz, F. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10): 534–559.

Supalla, T. 1986. The classifier system in American Sign Language. In *Noun Classification and Categorization* [Typological Studies in Language 7], C. Craig (ed.), 181–214. Amsterdam: John Benjamins.

Swabey, L. 2002. *The Cognitive Status, Form and Distribution of Referring Expressions in ASL and English Narratives*. PhD dissertation, University of Minnesota.

Wilbur, R. 1987. *American Sign Language: Linguistic and Applied Dimensions*. San Diego CA: College-Hill.

Zwitserlood, I. 2003. *Classifying Hand Configurations in Nederlandse Gebarentaal (Sign Language of the Netherlands)*. PhD dissertation, Utrecht University.

Zwitserlood, I. 2012. Classifiers. In *Sign Language. An International Handbook*, R. Pfau, M. Steinbach & B. Woll (eds), 158–186. Berlin: Mouton de Gruyter.

Zwitserlood, I. & van Gijn, I. 2006. Agreement phenomena in Sign Language of the Netherlands. In *Arguments and Agreement*, P. Ackema, P. Brandt, M. Schoorlemmer & F. Weermann (eds), 195–229. Oxford: OUP.

Between narrator and protagonist in fables of German Sign Language

Annika Herrmann and Nina-Kristin Pendzich

University of Hamburg / University of Göttingen

Sign languages have the unique property of simultaneously transmitting information in various ways. Within perspective taking and role shift, a signer may use various articulators such as the hands, the body, and the face to simultaneously represent different protagonists and/or the narrator. This paper discusses data from German Sign Language (DGS) with regard to parallel perspectivation in role shift, in particular action role shift found in the DGS data set of the Aesop fables. We categorize the different types of parallel perspectivation and investigate classifiers within action role shift as a phenomenon at the gesture-sign interface.

1. Introduction

In sign languages, perspective shift is an important and very frequent discourse structuring mechanism that signers use to take over the role of another referent or fictional character. Such a role shift is used for the reproduction of utterances and thoughts of another character (called *quotation role shift*, *constructed dialogue*, *constructed discourse,* etc.) and it may also be used for the reproduction of actions, emotional states, and mannerisms (called *action role shift, constructed action, role playing,* etc.). In this paper, we focus on the latter type. When investigating instances of action role shift, it is significant to consider simultaneity in sign languages. The layering of linguistic information in sign languages is generally achieved by the simultaneous use of different manual and nonmanual articulators such as the hands, the head, the face, and the torso. Obviously, apart from grammatical articulation, gestural means can also be layered on top of each other and appear parallel to signing.

Since several linguistic studies have dealt with perspective shift, we provide a brief overview of current analyses proposed for the different types of role shift and clarify the terminology that is used in this paper (Section 2). Our overall focus is

© 2018 John Benjamins Publishing Company

the analysis of parallel perspectivation within action role shift in fables of German Sign Language (DGS) by discussing the possibilities of layering and the interaction of the perspectives of protagonists and the narrator (Section 3). We provide an overview of the state of the art and discuss research approaches concerning first the application of upper body parts as linguistic devices for simultaneous perspectivations (Section 3.1.1) and second the use of classifiers (Section 3.1.2). Our study is based on a data set of nine DGS fables from three signers (Section 3.2). We consider different types of parallel perspectivation (Section 3.3) and particularly analyze the properties of simultaneous expressions performed by the narrator in role shifting (Section 3.4). Within action role shift, gestural acting can be implemented into narration. Role shift builds a grammatical frame in which gestural means can be layered with linguistic means. Thus, we consider action role shift a phenomenon at the gesture-grammar interface (Section 3.5). In order to approach a unified account for both quotation role shift and action role shift, we suggest an extended agreement analysis (Section 4). To conclude the paper, we provide a brief summary and point to specific issues for further research (Section 5).

2.   Role shift in sign languages

2.1   Terminology

In sign language discourse, the reproduction of words, thoughts, manual and non-manual gestures, emotional states, mannerisms, behaviors, events, and actions of real persons or fictional characters is exceedingly present. Indisputably, these diverse forms of reproduction are universal in the sense that they are used in all languages irrespective of the visual-gestural or auditory-oral modality, thus in signed and spoken languages (cf. Lillo-Martin 2012: 373). However, both modalities show differences with regard to the linguistic constraints, the integration in the language system and the form of implementation (see Sections 2.2 and 2.3). Concerning the reproduction of action in the broadest sense, further differences relate to the frequency of the phenomenon in spoken and sign languages (see Section 2.3). In the following, we use the term *role shift* as an umbrella term that includes the different types of shifts in sign languages. It is a modality-specific way of perspective shift in which the signer reproduces someone else's words or actions in a similar way to direct and indirect quotation in spoken languages. Role shift is particularly applied during narrations but by no means it is restricted to that use.

Regarding spoken languages, Clark and Gerrig (1990: 793) note that one difference between quotations and descriptions is that the former gives a *direct experience*.

> When we hear an event described, we interpret the speaker's words and imagine the event described. But when we hear an event quoted, it is as if we directly experience the depicted aspects of the original event. We perceive the depictive aspects partly as we would the aspects they are intended to depict.
>
> (Clark & Gerrig 1990: 793)

Equally, the use of role shift in sign languages enables a direct experience. Therby, it is save to say that, in general, two types of role shift are distinguished, even though the terminology for both phenomena is not consistent throughout the literature.[1] For our purposes, the terms *quotation role shift* (Q-RS) and *action role shift* (A-RS) make perfect sense. Q-RS means that a signer reproduces words, thoughts, or is retelling conversations between two or more referents. In contrast, A-RS refers to the reproduction of actions, emotional states, and mannerisms. Here, the use of manual and nonmanual gestures is particularly essential. One main difference to the terminology that Metzger (1995) and Cormier et al. (2011, 2015) use is the fact that we see *role shift* as an umbrella term whereas they define *constructed action* as the collective term. Consequently, they consider constructed dialogue as one type of constructed action. Lillo-Martin (2012: 376) accentuates, that "[t]he description of role shift as a type of constructed action recognizes that many components of this phenomenon are analogous to the use of gestures and changes in voice quality during narration in spoken languages." We do not deny the interplay between both types. Yet, in sign languages, we may observe the reporting of words or rather signs without parallel constructed action such as the gestural-affective imitation of a protagonist. Hence, we suggest to apply the more general term *role shift* as the basic category and distinguish between the different uses like pure quotation and pure imitation as two ends of a continuum. This is actually quite similar to the division quotational and non-quotational use of role shift by Pfau and Quer (2010), Lillo-Martin (2012), and Schlenker (2017). As mentioned by Tannen (2007[2]), Clark and Gerrig (1990) and others for spoken language, direct quotation does

---

1.   The different uses of a perspective shift have been given various names. Metzger (1995) and Cormier et al. (2011) take *constructed action* as the umbrella term and consider *constructed dialogue* (here *quotation role shift*) as a form of *constructed action* (here *action role shift)*. Engberg-Pedersen (1993, 1995) distinguishes between *shifted reference, shifted locus,* and *shifted attribution of expressive elements.* Lillo-Martin (1995, 2012) uses the terms *reference shift, shifting reference, utterance reports,* and *constructed action.* Quinto-Pozos (2007) notes that the signer is *becoming the object* and calls it *constructed action.* Pfau and Quer (2010) and Lillo-Martin (2012) oppose *quotational* with *non-quotational uses of role shift.* Herrmann and Steinbach (2010, 2012) adopt the same classification but have called the categories *role shift* and *constructed action.* Schlenker (2017) distinguishes between *attitude role shift* and *action role shift.* Further terms that can be found in this research area are *perspective shift, referential shift, role playing, role taking, role switching,* and *body shift.*

not neccessarily need to be a verbatim report of the originally spoken words but a *constructed dialogue* (Tannen 2007^2: 17, 111–112) and a *selective depiction* (Clark & Gerrig 1990: 795). Tannen (2007^2: 107f.) emphasizes "that 'reported speech' is not reported at all but is creatively constructed by a current speaker in a current situation." Likewise, quotations in sign languages are not straight copies of speech or action of a real person or fictional protagonist but rather constructions of the present signer (cf. Metzger 1995: 257; Liddell & Metzger 1998: 660). Within reported speech, utterances have been taken out of the original context to another context (cf. Tannen 2007^2: 108, 112). Clark and Gerrig (1990: 770) use the terms *current domain* and *source domain* and describe that each of these is specified by at least six parameters: "speaker, addressees, place, time, vantage point, and action." The same holds for role shift in sign languages.

It is important to note that there are obviously several overlappings between Q-RS and A-RS. In regular quotation role shift, on the one hand, signers may of course use affective facial expressions and role playing of the person or protagonist whose utterances or thoughts they reproduce. On the other hand, signers may quote within a phase of action role shift (cf. Pfau & Quer 2010: 397; Metzger 1995: 261). Hence, the relationship between these different types of shifting is much more complex than it seems at first sight and the interaction between both asks for a unified treatment. What unifies Q-RS and A-RS is that both represent an event from another character's point of view, as interpreted by the signer (cf. Lillo-Martin 2012: 370).

2.2 Quotation role shift

In case the signer reproduces actual words or utterances, we deal with a sign language specific means of quotation, which unites properties of both direct and indirect speech. A typical example of quotation role shift in DGS is provided in (1). When quoting what Tim said to Anna, the signer shifts into the perspective of Tim, who is the reported signer.

$$\overline{\hspace{4em}}^{\text{rs}}$$

(1) TIM IX$_{3a}$ ANNA IX$_{3b}$ SAY : $\overline{\text{TOMORROW }_1\text{HELP}_2}$ (DGS)
    'Tim said to Anna that he will help her tomorrow.'
    'Tim said to Anna: "Tomorrow I will help you."'

(Herrmann & Steinbach 2012: 211)

Signs are glossed in small caps, indices represent different loci in the signing space, and the colon stands for a prosodic pause. The specific context shift in Q-RS which refers to signers and addressees of a different utterance context is indicated by 'rs' and the line above the glosses, which illustrates the scope of the role shift

nonmanuals. The material below the line is interpreted as an embedded sentence that has formely been uttered by Tim (3a), the signer of a different context, not the actual context, and who is signing to Anna (3b). Both translations as indirect and direct speech are provided in the example to indicate that Q-RS is neither the former nor the latter, thus, we remain neutral with regard to the direct/indirect texture.

To clarify in detail what exactly triggers the context shift, it is necessary to investigate nonmanual markers and the interplay between them, as well as gestural components of signing. The formal features that mark role shift in general are body leans, head positions, eye gaze changes, and facial expressions (for ASL, see Winston 1991: 404). Herrmann and Steinbach (2012) argue that the latter do not relate to loci in signing space and are thus not part of the agreement like marking of Q-RS. Facial expressions have a different function within reported speech or reported action and are either lexical or grammatical markings depending on the signs and utterances or relate to the imitation of the quoted signer. The grammatical body leans, head positions and eye gaze changes to mark the shift are in itself not an imitation of the reported signer, but solely indicators of the context shift, not necessarily meaning that the person changed their body position, for instance. Thus, Herrmann & Steinbach (2012: 213–214, 220) look at the three markers of roles shift that occur in specific combinations of minimal and maximal markings.

The nonmanual indicators of Q-RS usually have scope over the entire quoted material and occur simultaneously to the manual signing stream. The features eye gaze and head position turn towards the right or left in order to mark the addressee of the reported utterance. Body lean indicates the shift to the reported signer. A prototypical and maximally marked quotation role shift involves all of these nonmanual markers, but it can also be marked by an eye gaze change alone, for instance. "The nonmanual marking is often so subtle that it escapes nonnative signers" (Quer 2011: 279).

In case the signer adopts affective and gestural facial expressions or a specific signing style of the reported signer, this can be compared to intonational variations of the quoted material to indicate the reported speaker's voice or his way of speaking in direct quotations in spoken languages (cf. Tannen 2007; Günthner 1999, 2002; Herrmann & Steinbach 2012). Thus, the imitation of characters to indicate direct quotation is not restricted to the visual-gestural modality, but can be found in spoken languages, too (see also Stec 2012 and Stec et al. 2016 for multi-modal role shift in spoken language narration). Nevertheless, quotation role shift is a modality-specific phenomenon as it uses grammaticalized nonmanuals in a systematic way to mark quoted utterances. Particularly with regard to action role shift, however, it is possible to indicate a perspective shift by gestural facial expressions and possibly manual gestures alone, so gestures may indeed be used to trigger

a context shift (cf. Lillo-Martin 2012; see Section 2.3). This is frequently found in situations with clear reference tracking and during the imitation of reported events in which action reports and speech reports are combined.

In the following, we briefly sketch the analyses of Q-RS in the literature. Since our focus is more on A-RS and the discussion of mixed perspectivation, this will not be an exhaustive discussion (for a detailed overview, see Lillo-Martin 2012). In most analyses presented so far, the context shift is triggered by some kind of operator in a higher domain. Lillo-Martin (1995, 2012: 376, 380) proposed a point of view (POV) predicate for ASL that takes a complement clause including a syntactic operator in the C-domain that binds indexical pronouns. She states that first-person pronouns in the complement of the POV predicate are logophoric and are interpreted as co-referential with the subject of the POV, rather than with the speaker. This is similar to analyses of languages such as Abe, Ewe, and Gokana, where logophoric first person pronouns may refer to the subject of a matrix-clause. Running into a few problems, such as the double occupation of the embedded C-domain when embedding interrogatives or sentences with topicalized constituents, the subsequent analyses try to avoid a bound syntactic operator. Quer (2011: 294–296) builds on basic assumptions of Lillo-Martin (1995), but says that Q-RS involves a covert point of view operator in the left periphery (*speech act phrase* following Speas & Tenny 2003) rather than a POV predicate. This operator materializes in the Q-RS nonmanual morphology and accounts for the shifted interpretation of indexicals in its scope. In particular, he extended the structure to include non-pronominal indexicals such as time and locative indexicals (Quer 2011: 294–296).

As briefly mentioned above, Herrmann and Steinbach (2012) suggest a nonmanual agreement analysis for role shift and – except for obvious differences such as the domain of agreement (verb vs. clause), for instance – they highlight the similarities between Q-RS and verb agreement. The nonmanuals agree with the loci of the signer and the addressee of the reported utterance. Instead of a purely semantic context shift operator, that materalizes in nonmanual morphology as in Quer's (2011) account, a nonmanual agreement operator in a high functional position of the clause triggers a context shift from the actual context to the context of the reported utterance (cf. Herrmann & Steinbach 2012: 203, 221; see Section 4.2 for an extension of the nonmanual agreement analysis to a unified account for Q-RS and A-RS). Thinking a bit further, we need to broaden our view of context shifts and also take reported actions into account. Role shift is a grammaticalized way of quotation that is restricted in many ways and systematically marked by specific features. However, even though A-RS mainly involves gestures, it is also restricted in certain ways by, for instance, the signing space and the use of specific types of classifiers. We will elaborate on these issues in the following section.

2.3   Action role shift

For the classification of constructed action, in our terms action role shift (A-RS), it is essential to look at the characteristics with respect to form and function. This specific type of shift is used to reproduce events, actions, manual and nonmanual gestures, emotional states, behaviors and mannerisms of another person, fictional character or personified object. Hence, A-RS is a *depiction* or a *demonstration* and not a *description* of another referent (cf. Clark & Gerrig 1990: 764–768).[2] As Clark and Gerrig (1990: 802) accentuate, people can use three devices in communication: "They can describe. They can indicate, or point. And they can demonstrate." By means of A-RS recipients have the opportunity to witness an action taking place directly in front of their eyes (cf. Liddell & Metzger 1998: 694). A-RS is expressed by the upper parts of the body that are involved in the reported event (cf. Metzger 1995: 262). The signer's direction of eye gaze, for instance, indicates the direction towards which somebody is actually looking (see also Liddell & Metzger 1998: 671–672; Janzen 2004: 163). This shows that A-RS can be triggered solely by nonmanual articulators, at the same time leaving the hands free for independent articulation (see Section 3). However, it is not only the imitation of facial expressions that is essential. A-RS can generally be defined as manual and/or nonmanual embodiment. This contrasting of characters by body shifts, changes of facial expressions and eye gaze has been found extensively across all sign languages (see Padden 1986 for an early reference). Note that in A-RS, the torso, the head, and the eye gaze can but do not need to be directed to the addressee of the reported context. Instead of the use of signs as in Q-RS, in this type of role shift we deal with the imitation of activities and the use of manual and nonmanual gestures. Lillo-Martin and Klima (1990) also note that a change in signing style could in principle indicate the perspective shift. A full understanding of the semantics of reported events is based on decoding both types of information: signed and gestured (cf. Liddell & Metzger 1998: 657–659). A-RS constitutes a dynamic interface between grammar and gesture. Thus, various terms can be found trying to define this interface nature of A-RS as embodiment, such as "body-quotation" (Lillo-Martin 2012: 367, 369f.).

Another relevant aspect is the relation between the occurence of A-RS and different types of text. Corresponding to previous research, the use of A-RS is hardly dependent on the type of text but highly dependent on the selected strategy of expression. If a signer puts an animate or personified entity in focus, the use of A-RS is decisive (cf. Fischer & Kollien 2010: 505, 508). One example provided by Fischer and Kollien (2010: 505–506) is an A-RS for the expression of *electrosmog*. They

---

2. In Clark and Gerrig's (1990) approach concerning spoken languages, quotations are treated as demonstrations.

outline that an informant performed a parallelized A-RS in which a non-designated, animate entity is used in order to look to a cloud. Fischer and Kollien (2010: 506) classify this A-RS as 'weakly pronounced'. Examples like this show that, apart from the use of A-RS in signed narration, A-RS also occurs to a significant extent in non-narrative texts (cf. Fischer & Kollien 2010: 504, 508). A typical example of a simple A-RS in DGS is given in (2) and involves gestural and affective imitations of a fictional character in the fable *The shepherd's boy and the wolf*. Prior to the A-RS meaning 'the boy stands holding a stick, looking around, being bored and irritated', the context, paraphrased as *It was nice everywhere, but boring and always the same*, is articulated with regular signs.

$$\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{\text{rs}}$$

(2) *stand-hold-a-stick-look-around-be-bored-and-irritated*

The A-RS in (2) illustrated in Figure 1 is used to shift into the perspective of the shepherd's boy by exemplifying the boring repetition of his actions with additional affective facial expressions of being bored. The example shows that for an A-RS, a modification of the upper body position to the previously established loci in the signing space is not required. In line with this, Janzen (2004: 152–153) notes that a signer may use a *mental rotation of the space* instead of a *physical shift in space* for taking another perspective. Janzen (2004: 162) displays an example in which a "signer does not move around the space; the space rotates and no physical shifts into various loci in the space take place." Within *mental spatial rotation,* the coding of perspective is more implicit and requires more inference by the addressee (Janzen 2004; Graumann 2002). This fits with the observation that in general, learners of a sign language as second language cannot effortlessly handle A-RS (cf. Lillo-Martin 2012: 375; Metzger 1995: 263; von Randow 2016). In (2), the space of the narrator becomes the space of the shepherd's boy (cf. Janzen 2004: 158).

stand-hold-a-stick-look-around-be-bored-and-irritated

Figure 1. Simple A-RS (*The shepherd's boy and the wolf*)

Another instance of A-RS is illustrated in (3) which is a short excerpt from the fable *The tortoise and the hare* signed in DGS. The A-RS$_{hare}$ is integrated in a larger

Q-RS$_{hare}$ in which the hare is talking to the tortoise, laughing, and making jokes about her low speed. This mocking is illustrated by the embedded A-RS for the slow pace of the tortoise.

$$\overline{\phantom{\text{... GO imitate-slow-pace-of-the-tortoise SLOW GO}}}^{\text{rs}}$$

(3)　… GO *imitate-slow-pace-of-the-tortoise* SLOW GO

Prior to the A-RS$_{hare}$, the body of the signer is clearly shifted to the left, the established locus for the hare, and the signer looks to the right, the locus of the tortoise as the addressee of the Q-RS$_{hare}$. The first picture in Figure 2 shows the sign GO, followed by the A-RS$_{hare}$ illustrating the same action in more detail. During the imitation of the tortoise, the body is oriented to the center. The A-RS$_{hare}$ is followed by the signs SLOW and GO that frame the imitation of the tortoise. Following the A-RS$_{hare}$, the body is progressively shifted to the left side again, indicating that the signer is back in the Q-RS$_{hare}$ mode.

| GO | *imitate-slow-pace-of-the-tortoise* | SLOW | GO |
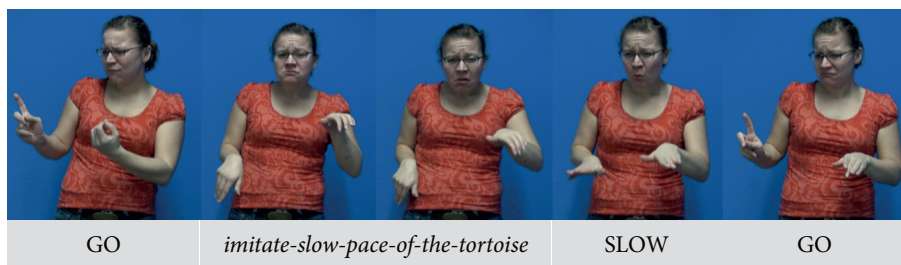
Figure 2.　Integration of simple A-RS in Q-RS (*The tortoise and the hare*)

Another interesting point is that the hands of the signer stand for the limbs of the tortoise in the A-RS$_{hare}$ in Figure 2. Here, we deal with the interaction between A-RS and grammatical classifiers, i.e., the grammar-gesture-interface on which we focus in Section 3.2.1 and 3.3. Both examples of an A-RS, so far, are cases in which the fictional protagonist is the agent. As Janzen (2004: 167) states, A-RS can also be used to denote which perspectives are excluded. If the signer turns his head explicitly aside, this may create the impression of not observing the simultaneously signed event. "The effect is one of nonparticipation in the event" (Janzen 2004: 167). By this gestural head turn "the narrator introduces a new character, however peripherally, and chooses to say something about this entity's (non)involvement in the event" (Janzen 2004: 167).

　　More subtle cases of A-RS, such as affective facial expressions of another person or fictional character which a signer may combine with Q-RS, are comparable with instances in spoken languages where a speaker uses the voice quality to indicate different referents in a reported dialogue (cf. Lillo-Martin 2012: 371; Liddell 1998;

Engberg-Pedersen 1993, 1995; Clark & Gerrig 1990: 775–777). Also, Liddell and Metzger (1998: 675) point to similarities between A-RS in sign languages and *shifted situations* in spoken languages. Parallel to words, speakers use different gestures to shift into a role of another character and to transmit his actions, words, or thoughts. These gestures, on the one hand, are *vocal gestures* including elements such as loudness, pitch, tempo, and accent (cf. Günthner 1999) and on the other hand *visible gestures* comprising nonmanual gestures (facial expressions, eye gaze, posture) and manual gestures (cf. Liddell 1998). Thereby, the only main difference is a modality issue. In sign languages, all of the different forms of information transfer are articulated by movements of the body, the hands, and facial muscles. Thus, the distinction between gestural and grammatical elements is more complicated (cf. Lillo-Martin 2012: 374). In spoken languages, speech may be complemented with pantomime, for which Clark and Gerrig (1990: 764) apply the term *demonstrate,* in the meaning of 'illustrate by exemplification'. They describe that "[d]emonstrations work by enabling others to experience what it is like to perceive the things depicted" (Clark & Gerrig 1990: 765). This function also applies to A-RS but it is really important to note the differences regarding the form. In contrast to pantomime in spoken languages, A-RS in sign languages is restricted to the upper parts of the body. For example, if a speaker wants to mimetically illustrate in the oral-auditory modality that a person walks on the spot it may be realised by using the legs. However, in sign languages, this is not expressed likewise even within A-RS because the legs are no linguistic articulator. Instead, additional means such as classifier conctructions are used (see Section 3.1.2 and 3.3).

Fischer and Kollien (2010: 506) compare the frequency of occurrence of A-RS used in meaning explanations given in DGS with specific gestures used in meaning explanations given in German (e.g. pointing gestures and enactment). For the latter, their results show a lower frequeny of occurrence. The rare instances of enactment as a manual or bodily demonstration of an action in German is particularly interesting. Fischer and Kollien (2010: 507–508) note that a pronounced role adoption with expressive facial expressions or body movements was almost absent in the elicited data of German speakers. They interpret this result as a consequence of the informant's choice for a formal register in the explanation of meanings. Similarly, Quinto-Pozos and Mehta (2010) note that the use of A-RS in sign languages varies depending on the addressee and also the formal or informal communication setting. Earis and Cormier (2013) compare the story telling of a fable in both BSL signers and English speakers. They note that for the same story, signers tend to use A-RS and the character's perspective more frequently than speakers, who prefer the narrator's perspective. A big discussion is also circulating around the question of the obligatoriness of A-RS in sign languages, as it is such an integral part of the language (cf. Quinto-Pozos 2007).

Due to the systematic integration of A-RS in sign languages, we do not find the identical frequency of such comparable strategies in spoken language. The complex interplay and the continuous shifts in the signing stream between narrator's comments, Q-RS and A-RS, are prominent properties of signed narration (cf. Lillo-Martin 2012: 373). It is interesting to investigate the two phenomena Q-RS and A-RS together, look at their interaction, combination, and work towards a unified analysis.

3.   Parallel perspectivation within action role shift in fables of German Sign Language

3.1   State of the art

Sign languages are articulated in the three-dimensional space "that in numerous ways mirrors the signer's conceptualization of some real space in which entities are situated and interact in some relative configuration" (Janzen 2004: 151). Thereby, a signer can choose between different kinds of perspectivation that lead to various possible representations of one event. Looking at simultaneous perspectivation in A-RS exhibits explicitly that A-RS is a highly complex linguistic phenomenon in sign languages. There are already several studies about parallel perspectives in different sign languages and diverse terms and approaches are used for similar phenomena. Aarons and Morgan (2003) investigate these issues for South African Sign Language (SASL), Metzger (1995), Liddell and Metzger (1998) and Dudis (2004) look at parallel perspectives in American Sign Language (ASL), Fischer and Kollien (2010) investigate DGS with regard to A-RS and also mention complex A-RS with multiple characters. Cormier et al. (2015) analyze A-RS in detail and discuss various criteria and the annotation of different types and degrees of A-RS. They note a degree-continuum of the intensity based on the number of active articulators used.[3]

One clear case of parallel perspectivation in ASL is Metzger's (1995: 263) example of *fist hit chin*. In this A-RS the fist of the signer represents the action of one character whereas the head constitutes another character's action. In the following,

---

3.   Fully agreeing with a thorough annotation and cateogorization of various types of roles including subtle A-RS, we would still like to carefully note that the amount of active articulators may just be due to the type of referent that is imitated or the type of behaviour of that referent. Intensity is an important and difficult measure, but a primary or secondary role should not be predicted based on the absolute number of active articulators. Furthermore, all articulators listed for A-RS may be used for lexical or grammatical purposes, so these markers need to be checked and eventually be excluded as they may not indicate an active role in those cases.

as basis for the discussion of the results of our study, we outline two fundamentals for perspectivations within A-RS: the upper body parts as linguistic devices (Section 3.1.1) and the use of classifiers (Section 3.1.2).

3.1.1   *Upper body parts as linguistic devices for simultaneous perspectivations*

Sign languages are expressed with manual and nonmanual articulators: the two hands, the torso, the head, and the face. The face can be further split up into smaller articulatory components such as the eyes, the eyebrows, the nose, the cheeks, the mouth, and the tongue. Via these manual and nonmanual channels linguistic information can be conveyed simultaneously. Furthermore, in sign language grammar, the strategy called *body as subject* (Meir et al. 2007) is of utmost importance. To investigate this strategy, Meir et al. (2007) analyze iconic or partly iconic *body-anchored verbs* such as the sign EAT in Israeli Sign Language (Israeli SL) and ASL:

> [T]he body, constituting one of the formational components of the sign, represents one particular argument in the event, the agent. It is important to note that the body does not represent first person. The sign EAT is signed on the mouth of the signer whether the subject in a particular event of eating is first, second or third person. In other words, the sign EAT has one form in all three sentences 'I eat' 'you eat' or 's/he eats' and this form is signed on the signer's mouth.   (Meir et al. 2007: 543)

Meir et al. (2007: 544) conclude that 'body as subject' is the crucial lexicalization pattern for the representation of events or actions in sign languages. In addition, they address *mimetic depiction* whereby a "signer uses his/her body to express an event mimetically" (Meir et al. 2007: 548). In general, signers have two options for the reproduction of actions: On the one hand, they can use their hands to describe events such as, for instance, somebody or something falls down on the ground. On the other hand, they can articulate such actions with their body instead of their hands. Meir et al. (2007: 548) note for Israeli SL and ASL that "the body represents an event predicated of an animate subject, not an inanimate one." Moreover, they explain that the use of the body for inanimate objects occurs mostly for poetic and theatrical functions in the term of personification as objects adopt animate-like conditions (Meir et al. 2007: 548).

> Although the body is used as a formational element in the structure of signs, it is still a human body that 'belongs' to the signer, an animate being. As such, it is not used to stand for inanimate entities. In this sense then, we see persistence of properties of the human body in certain aspects of the linguistic behavior of the body.
> (Meir et al. 2007: 549)

Since nonmanual parts of the body can be grammatically used for the reproduction of actions, it becomes clear that the hands may simultaneously fulfill different functions. The independence of different articulators is essential for the expression of different perspectives at the same time. Dudis (2004: 227, 229, 235) describes *four partitionable zones of the body*: the two hands, the oral articulators, and facial expressions. Our data yields that for DGS there is at least one additional relevant partionable zone of the body: the nose, as it can be used to indicate an referent independently of the facial expressions. This means that various smaller parts of the body may represent different referents. In principle, this should be possible with ears, the mouth, hair, etc., which may be used in accordance with the strategy *body as subject* (Meir et al. 2007; see Section 3.4). Similarly to Fischer and Kollien (2010: 504f.), we distinguish between (i) simple, pure A-RS, that is the signer's body embodies solely one fictional character, another person, one personified object or one event, and (ii) complex A-RS in which by means of *partitionable zones of the body* multiple perspectives or event parts are expressed simultaneously.

In addition, for signed narration it is crucial that we scrutinize *blended spaces*[4] (cf. Liddell & Metzger 1998: 665–673). Here, the narrative space is opposed to the real space of the current signer and these *two input spaces* may in parts be mapped onto the present, blended space. For instance, a character may be mapped onto the torso, the head, and the facial expressions of the signer, whereas the hands articulate the narration of the action with signs (cf. Liddell & Metzger 1998: 666, 668). Such simultaneous combinations of the narrator's signing and A-RS in DGS are the main focus of the present study.

3.1.2   *Classifiers*

Sign languages exhibit a specific way of classifying objects by modification of signs. Thus, the elements for modification are generally referred to as classifiers that have a non-specific meaning.[5] In verbal classifiers, for instance, verbs are morphologically modified according to the properties of their arguments by a change in handshape to indicate salient properties. Distinctions of various classifiers as initially drawn by Supalla (1986) devide classifiers according to semantic criteria and many following

---

4. Liddell and Metzger (1998: 662f.) refer back to Fauconnier and Turner (1996) who "describe a general cognitive process which they refer to as blending. It is a process that operates over two mental spaces as inputs. Structure from the two input spaces is projected to a third space, which they refer to as the blend." Liddell (1995) discusses the interaction of *real space* and *surrogate space* and Dudis (2004) uses the term *real-space blends*.

5. There is much debate about whether these modifications should be called *classifiers* or *depicting handshapes*, for instance. We will use the term *classifiers* and refer the reader to Zwitserlood (2012) for an overview and a summary of the discussion in the literature.

researchers suggested various different criteria throughout the years (cf. Schick 1990; Aronoff et al. 2003; Emmorey 2003; Zwitserlood 2003; Benedicto & Brentari 2004; Schembri et al. 2005; Eccarius & Brentari 2007; Grose et al. 2007; Morgan & Woll 2003, 2007; Perniss 2007a, b).

As Zwitserlood (2012) points out, two major categories can be generally distinguished according to their functions: *whole entity classifiers* (also called *subject classifiers*) and *handling classifiers* (also called *object classifiers*). The former refer to handshape modifications that directly represent entities and objects while the latter specify the way of handling an object with the hands. In addition, signs called *size and shape specifiers* are a particular subset that is often included in the category of classifiers. They, however, may have nominal, adjectival or (ad)verbial functions and cannot combine with verbs. As they are a separate group and did not appear in our A-RS data, we will exclude them from our analysis and only look at classifiers that are morphological modifications of verbs. In A-RS, classifiers are used in cases where the limited signing space puts restrictions on the gestural imitation of actions. For instance, if the legs are involved in a specific action that is expressed by A-RS, the signer uses whole entity classifiers to avoid moving her actual limbs. This would be allowed in pantomime but the legs are not linguistic articulators in sign languages. Interesting cases occur when a handling classifier appears in A-RS, because the distinction between the real gestural action and the choice of a specific handshape is hard to test. The linguistic elements incorporated in A-RS reveal the systematic constraints that A-RS is subject to and provide an interesting case of multiple perspectivization, as the narrator's perspective and the grammatical elements are intervening in the gestural role play. Metzger (1995: 258, 265) also mentions the co-occurrence of A-RS with classifiers and considers them to be the reason for the vast diversity of forms in A-RS, but classifiers in A-RS are not her main point. Particulary interesting for our data is the work by Aarons and Morgan (2003) for South African Sign Language, as they also describe the use of classifier predicates in the creation of multiple perspectives. They provide an example of a parachutist's fear in A-RS while signing the classifier predicate FLY. In this paper, we are interested in a thorough description of how the interaction between narrator's perspective and protagonist's perspective works.

3.2   Methodology and data

The study is based on a data set of five Aesop's fables[6] elicited with three deaf DGS signers, age 28, 38, and 44. The youngest informant is a native signer in third generation and the two other informants are near-native signers.[7] All of them are members of the Deaf community. Whereas informant A and B are right-handed signers, informant C is a predominantly left-handed signer. Table 1 summarizes the metadata of the informants.

Table 1. Metadata of the informants

Informant	Age	Gender	Born deaf/ Deaf at the age of	Signing at the age of	Deaf or signing parents	First or preferred language	Dominant hand
A	44	f	No/3–4	6	No	DGS	right
B	38	m	No/2	3–4	No	DGS	right
C	28	f	Yes	0–1	Yes	DGS	left

The videos were recorded in the Experimental Sign Lab of the Georg-August-University in Göttingen and the elicitation procedure comprised two steps for each fable: (i) the informant was instructed to read and memorize a text version of one fable without time pressure, (ii) the informant then independently signed his own version of the fable without any interruption by the instructor. During the elicitation, one camcorder captured the signer's upper part of the body and a second camcorder zoomed in the signer's face. The videos were coded with the linguistic annotation software ELAN.[8] So far, we particularly investigated the form and function of simultaneous perspectives including the narrator and a protagonist. For the analysis presented in this paper, we evaluated a selected data set of nine DGS fables comprising the first three fables signed by each informant: *The sheperd's boy and the wolf, The tortoise and the hare* and *The lion and the mouse*. Each signed fable has a duration of around two minutes. Altogether we analyzed 17 minutes of signing.

---

6.   We recorded five of the Aesop's fables as used in the ECHO project. For further information, see Crasborn et al. (2007) and <http://www.sign-lang.ruhosting.nl/echo/>.

7.   Near-native status was defined as early sign language acquisition before the age of six. According to Morgan and Woll (2002), sign language performance and perceptional competences are best in situations of early acquisition before the age of seven, as this age limit shows effects on production and perception. Emmorey et al. (1995) define early signers by acquisition between 2–7 and late signers by acquisition from 10 on (also see Keyser & Larson-Hall 2005; Morford 2004).

8.   ELAN (European Distributed Corpora Project Linguistic Annotator) is an annotation tool particularly for multi-modal data developed at the Max Planck Institute for Psycholinguistics in Nijmengen. The program can be downloaded from <http://tla.mpi.nl/tools/tla-tools/elan/>.

3.3 Types of parallel perspectivation

When analyzing perspectivation in signed narration it is quite striking how fast the mapping of single articulators onto specific perspectives may change and switch. Figure 3 outlines the separate occurence of the narrator's speech, Q-RS, and A-RS as well as their complex interaction and layering. This reveals six different scenarios.

Scenario	I	II	III	IV	V	VI
Narrator	✔			✔	✔	
Q-RS		✔		✔		✔
A-RS			✔		✔	✔

Figure 3. Layering of narrator's perspective, Q-RS, and A-RS

In signed narration, Q-RS and A-RS are frequently used as strategies of story telling. There is a complex interaction between both and we find various instances of layering of these two types of role shift. Furthermore, these systematic shifts can be interrupted and accompanied by elements assigned to the narrator. Hence, the distinction between the perspectives of protagonists and the narrator is not always clear-cut and we deal with perspectives realized in different degrees. In general, this is not crucially different from spoken languages. "Speakers also combine words, gestures, facial expressions, and changes in voice quality to convey the same range of narrative components" (Lillo-Martin: 2012: 373).

For the realization of different perspectives, sign languages display two ways: (i) fast switching between perspectives, (ii) simultaneous layering of perspectives by the unique possibility of the parallel use of the different manual and nonmanual articulators. With regard to type (i), our data reveals fast switching between the narrator's perspective and the protagonist's perspective as well as between different protagonists. This is implemented by hand dominance shifts or by changing the nonmanual markers within a single utterance. This switching obviously may occur without explicitly mentioning that perspectives have shifted. (e.g., sequence of A-RS, signing from the narrator's point of view, and again A-RS). Nonmanual markers for switching between perspectives may be body shifts, head tilts, eye gaze changes, facial expressions, and mouthings. The latter means that during fast switching between perspectives, which can be quite subtle, the current perspective may be specified simultaneously by a mouthing that, for instance, reactivates a referent (see Figure 4).

| *detect-the-tortoise* | IX$_{3a}$ | SHIT | *walk-slowly* |
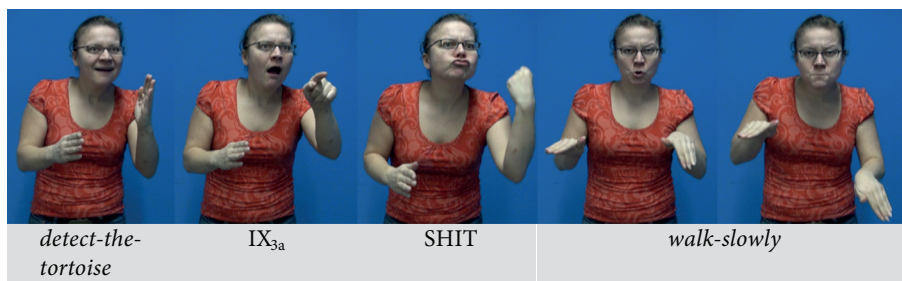
Figure 4. Fast switching between perspectives and reactivating reference to tortoise by mouthing (*The tortoise and the hare*)

$$\overline{}^{\text{rs}} \quad \overline{}^{/\text{tortoise}/ \text{ rs}}$$

(4)   *detect-the-tortoise* IX$_{3a}$ SHIT *walk-slowly*

In (4), taken from the fable *The tortoise and the hare,* the first A-RS from the hare's point of view and the second A-RS from the tortoise's point of view are clearly distinguishable through facial expressions, body shift, and eye gaze. The first A-RS$_{hare}$ portrays that the hare detects the tortoise in front of him. It is combined with an index targeted at the tortoise and followed by the sign SHIT. The subsequent A-RS$_{tortoise}$ represents the persistent slow walking of the tortoise. In addition to the explicit nonmanual indicators, in this extreme fast switch of perspectives the signer articulates the mouthing /*tortoise*/ from the narrator's point of view for reactivating the reference. While this short example already exhibits the parallel occurence of a character's perspective with some information given by the narrator, there are many cases in signed narration where more complex layerings are used. As pointed out before, different body parts may thus act independently showing distinct perspectives at the same time. In Table 2, we list the various articulatory combinations of such simultaneous layerings.

Table 2. Articulatory types of simultaneous layerings (perspective of protagonist and narrator)[*]

		Protagonist	Narrator
Simultaneous articulators	(a)	hand1 + body + face	hand2
	(b)	body + face	hand(s)
	(c)	body	hand(s) + face
	(d)	hand2 + body	hand1 + face
	(e)	hand(s) + body + face	mouth (for mouthing)

* Here, we only listed possible cases of two perspectives. Nevertheless, the combination of three perspectives is equally possible as later shown in further examples.

With these five different articulatory combinations of *partitionable zones of the body* (cf. Dudis 2004: 227) found in our data, the simultaneous embodiment of different perspectives is realized. Possibly, this is not an exhaustive list and the analysis of further signed narrations may result in additional combinations. In type (c), for instance, the body may take the role of a protagonist, while the face, specifically eye gaze, may indicate that the actual signing with the hand(s) belongs to the narrator. In the present study, we focus on the layering of the narrator's perspective and the protagonist's perspective by simultaneous articulations within A-RS. Obviously, further investigations should consider more closely which of the form types listed in Table 2 may also be used for the layering of two different protagonist's perspectives.

As mentioned above, we have evaluated three different fables *The shepherd's boy and the wolf, The tortoise and the hare,* and *The lion and the mouse,* each signed by three deaf DGS signers. Our counting method of A-RS is apriori similar to the following method provided by Metzger (1995: 261):

> From the time a character begins an action or series of actions to the end of that uninterrupted sequence is considered to be one occurrence of constructed action. Moreover, when the action shifts from one character to another, this is considered to be two occurences of constructed action.

However, during the analysis, further determinations in two directions were required. (i) If a single string of the plot composed of immediately connected actions is performed, it is counted as one A-RS like in the counting mechanism of Metzger (1995). But we found that, if in a series of actions, semantically and prosodically clearly distinctive action units occur, it is adequate to count them as several individual A-RSs. In the fable *The lion and the mouse*, for instance, an informant describes the scene in which the lion is caught in a trap of the hunter and is rescued by the mouse by using a sequence of different A-RSs. Here, one $A\text{-}RS_{lion}$ embodies the 'hanging in the net' and the subsequent new $A\text{-}RS_{lion}$ illustrates the 'step by step falling out of the net' (ii) Of particular interest for the present study are the occurences of A-RS combined with a parallel narrator's addition. The simultaneity of both perspectives is indicated by the perseveration of one hand while signing one-handed signs, by keeping the nonmanual embodiment while using one/two-handed signs or by a combination of these manual and nonmanual means. A special case is the simultaneous use of a narrator's mouthing during A-RS (see Table 2). The whole unit is counted as one A-RS including a narrator's addition (see Section 3.4). Hence, we make a categorical distinction between simple A-RS and simultaneous constructions of A-RS + narrator's comments.

In our data, we found 183 A-RSs in total (see Table 3). Out of these, 121 are simple A-RSs and 62 are simultaneous combinations of A-RS + narrator. Hence, a really interesting result is that in 33,9% of all A-RSs, the current narrator context is

activated simultaneously. This enhances the idea that concurrent perspectivation is an essential strategy of story telling in sign languages and statistically proves it for DGS.

Table 3. Statistics for simultaneous perspectivation in fables of DGS

Fable/Informant	A-RS simple	A-RS + narrator simultaneous	A-RS total
1/A	14 = 70%	6 = 30%	20
1/B	15 = 60%	10 = 40%	25
1/C	11 = 64,7%	6 = 35,3%	17
2/A	13 = 76,5%	4 = 23,5%	17
2/B	13 = 61,9%	8 = 38,1%	21
2/C	23 = 82, 1%	5 = 17,9%	28
3/A	8 = 66,7%	4 = 33,3%	12
3/B	8 = 61,5%	5 = 38,5%	13
3/C	16 = 53,3%	14 = 46,7%	30
Total	121 = 66,1%	62 = 33,9%	183

For the theoretical classification of simultaneous perspectivation including A-RSs and narrator's additions, it is fruitful to compare these mixed perspectivations in signed action quotes with *hybrid quotations* as defined by Clark and Gerrig (1990) for spoken languages. They note that "the essence of quotation is demonstrating something rather than describing it, and demonstrations come mixed with descriptions in hybrids of many forms" (Clark & Gerrig 1990: 791). One case of such hybrid quotations is the use of indirect quotation combined with *concurrent demonstrations* that may be expressed as intonational changes. Clark and Gerrig (1990: 791) illustrate this with the following example:

(5)   And then Mrs. Dewlap said that he [raising voice] COULD JUST WAIT FOR HIS TURN WITH THE REST OF THEM [lowering voice] and so he did.

In this case, the speaker demonstrates the yelling of the woman by raising the voice although the speaker is describing her utterance with an indirect quotation (cf. Clark & Gerrig 1990: 791). "[T]here is no sharp division between expressions that depict and those that describe. Some do both" (Clark & Gerrig 1990: 792). The same holds for A-RS in sign languages even though some other characteristics are striking as well. As opposed to Q-RS, in A-RS, actions are reproduced instead of speech, which is in itself different to Example (5). In A-RS, the demonstration is the underlying basis and the description is added. However, both are used simultaneously, which is the central point. In this respect, classifier constructions, which we discussed in Section 3.1.2, are relevant with regard to the hybrid nature of mixed A-RSs. The use of classifiers in A-RSs can be understood as a combination of

depiction, i.e., demonstration (A-RS), and description (classifiers). Pursueing this approach, it seems appropriate to capture cases of A-RS layered with an addition from the narrator's point of view as hybrid quotation. Accordingly, the action is depictively reproduced and combined with descriptive elements from the narrator (see Section 4 for the integration of this approach in the nonmanual agreement analysis).

Another form of such a mixed quotation discussed in the spoken language literature is Free Indirect Discourse (FID). However, one main difference between FID and role shift is that the former occurs mainly in fictional narratives and is not frequent in face-to-face communication (for the latter, see Clark & Gerrig 1990: 787). Both FID and Q-RS share some properties that allow for a fruitful comparison. However, as FID is currently compared to Q-RS alone, we will not go into detail here, but refer the reader to Maier (2015) and Hübl (2013, 2014) and the discussion on conventionalized but pragmatic principles of unquotation.

3.4   Simultaneously layered additions by the narrator within action role shift

In spoken and signed narration, many shifts between the perspectives of narrator and character occur (cf. also Janzen 2004: 152). However, it should be noted that a shift to the perspective of the narrator is not necessarily always a shift to neutral space since also the space of the narrator may be oriented towards a specific perspective (Janzen 2004: 170).[9] In order to be classified as an addition of the narrator, the respective sign(s) must clearly not be interpreted as a comment or thought of a protagonist whose action is reproduced but as a narrator's addition to reinforce or complete the concurrent A-RS (cf. also Metzger 1995: 264). The narrator's addition (i) may be indicated by a change in eye gaze to the addressee of the actual context or (ii) may be so deeply incorporated in the action unit that no change in eye gaze occurs. Regarding (i), it has to be mentioned that eye gaze can be directed to the addressee either during the whole addition or it is used only briefly to indicate the addition. Concerning (ii), in some cases, an absent change in eye gaze results from the fact that the eye gaze is already oriented towards the actual addressee during the A-RS. An example for this type, in which the eye gaze directed to the front is maintained during the addition of the narrator, is given in Figure 5.

---

9.   Janzen (2004: 170–171) concentrates on "shifts to third-person singular story character perspectives" and notes that "features that distinguish these from narrator perspective deserve future attention."

Figure 5. A-RS with a simultaneous narrator's addition, but no eye gaze change (*The tortoise and the hare*)

<div>

(6) **Protagonist** ^{rs} *run-with-increasing-exhaustion*
     Narrator   EXHAUSTING----|

</div>

In Figure 5, the first picture illustrates an A-RS$_{hare}$ with the meaning 'running with increasing exhaustion.' The subsequent pictures show the simultaneous addition of the narrator in the form of the descriptive sign EXHAUSTING. The sign does not interrupt the continuous facial expressions and posture of the A-RS$_{hare}$ and can be analyzed as directly co-occuring with the A-RS$_{hare}$. Immediately afterwards, another parallel construction (A-RS$_{hare}$ + narrator) follows and represents 'being exhausted and lying down.'

Focusing on the functional aspects of concurrent additions by the narrator in A-RS, the data yields the following four main categories (see Table 4).

Table 4. Functions of simultaneous additions of the narrator in A-RS

(a) Description (signed explanation of action, cf. Metzger 1995)
(b) Reactivating reference
(c) Adding supplementary information
     – Supplementary information about the same protagonist
     – Supplementary information about a different protagonist
     – General supplementary information
(d) Comment of the narrator's attitude

Following Metzger (1995: 262), we use the term *description* for cases in which the signer adds "a form of narration along with the action constructions as an indirect description of events." These minimal signed descriptions of the simultaneously visible action of a character mostly consist of one single sign (cf. Metzger 1995: 263, *minimal narration, minimal comment*). An example by Metzger (1995: 264) is the usage of the sign LOOK-UP as a simultaneous indirect description of the A-RS 'looks up and to the left.' In the majority of cases, the descriptions of concurrent actions of

a protagonist are applied because the actions may otherwise not be visible enough. Another such an example by Metzger (1995) is CHEW. Using simultaneous signs to more clearly express the content of the A-RS is convenient in sign languages and strengthens the propositional content of the utterance. Here, a parallel to spoken languages can be drawn. Clark and Gerrig (1990: 766) explain that *demonstrations* in spoken languages have two different functions and illustrate them with the example *limp*. First, a demonstration can be an embedded *component part* of speaking. In this case, it may be introduced with a sentence such as 'The man walks like this' and after a demonstration of the man's limping the speaking is resumed. Second, and at this point relevant for our observations in DGS, a demonstration can be a *concurrent part* of speaking. This means while saying something like 'The man walked suddenly different,' the speaker might actually imitate the limping simultaneously (Clark & Gerrig 1990: 766–767). Such concurrent demonstrations occur in various types, comprising the use of tonal changes of the voice and facial expressions, among other markers (cf. Clark & Gerrig 1990: 791). A further very interesting analogy can be found with specific *sound quotations* or onomatopoetica in spoken languages. Clark and Gerrig (1990: 788–789) explain that these sound quotations may contain a descriptive part. Among other examples, they exemplify this with the expression 'knock, knock, knock, knock.' If these words are spoken very rhythmically in a slow and loud manner, then specific characteristics of a sound or action are depicted by the speaker. However, at the same time this "depiction is carried by a verb that describes the action as well" (Clark & Gerrig 1990: 789). In 'knock, knock, knock, knock,' the words are descriptive but timing, rhythm, and loudness may be *depictive* (cf. Clark & Gerrig 1990: 789).

Two further examples for *description* in sign languages are the already discussed combinations of A-RS$_{hare}$ + narrator in (6) and the A-RS$_{shepherd's\ boy}$ + narrator in (7) which frames a Q-RS.

		rs
(7) **Protagonist** *think-hold-chin scream--*|	HEY WOLF IX$_{3a}$ SHEEP EAT IX$_{3a}$HEY *scream----*|	
Narrator	SCREAM	SCREAM++

The excerpt of the fable *The shepherd's boy and the wolf* in (7) exemplifies a role shift that is composed of reported action and reported speech. Simultaneous *description* by the narrator occurs in both A-RSs (see the second and the last picture in Figure 6). The first picture in Figure 6 illustrates an A-RS$_{boy}$ realized with the entire body including facial expressions, posture, and the hands. In the following, the A-RS$_{boy}$ is still expressed by facial expressions and posture but the hands are not used for this perspective. The sign SCREAM is added by the narrator as a signed description of the simultaneously visible action of the boy's screaming. The use of the hands is not required for the nonmanual A-RS$_{boy}$, but the addition guarantees

that the recipient clearly understands the specific meaning of the action unit. A signed description by the narrator clarifies the action of the A-RS which otherwise would be more vague. This strategy can be found quite frequently in our data. In the glosses in (7), the small dashes indicate the articulation length of the signed description. It is striking that the articulation of the sign and the A-RS$_{boy}$, both expressing the meaning *scream,* are exactly coordinated in terms of length. In addition, the A-RS$_{boy}$ functions as an introduction to the Q-RS$_{boy}$ HEY WOLF IX$_{3a}$ SHEEP EAT IX$_{3a}$ HEY and replaces a matrix clause. During the signed description SCREAM, eye gaze is not directed to the addressee of the actual context but it is obvious that semantically the sign SCREAM cannot be a part of the Q-RS$_{boy}$ (the third picture in Figure 6 shows the sign WOLF as one element of the Q-RS$_{boy}$). Hence, the sign SCREAM is given from the perspective of the narrator. Furthermore, the same A-RS$_{boy}$ with the signed description SCREAM re-occurs after the Q-RS$_{boy}$ as illustrated in the last picture. Thus, the A-RS frames the Q-RS. As shown in Figure 7 illustrating the example (8), another informant performs a similar A-RS + narrator construction for the screaming of the boy.

| *think-hold-chin* | *scream* SCREAM | WOLF | *scream* SCREAM |

Figure 6. A-RS is used to frame Q-RS and contains a parallel description by the narrator (*The shepherd's boy and the wolf*)

| *have-an-idea* | WOLF | *scream* SCREAM |

Figure 7. A-RS with layered description by the narrator (*The shepherd's boy and the wolf*)

(8) **Protagonist** *have-an-idea* WOLF THERE WOLF THERE *scream--|*
   rs
Narrator SCREAM

The first picture in Figure 7 presents a simple A-RS that is used for the shift into the perspective of the boy: 'he is standing and has an idea.' This A-RS$_{boy}$ merges directly into the Q-RS$_{boy}$ that starts with the sign WOLF. Again, the A-RS$_{boy}$ is expressed with the face and the body and is accompanied by the sign SCREAM from the point of view of the narrator as a description of the parallel action. The nonmanual embodiment by facial expression is depicted in the last picture.

The second function of simultaneous narrator's additions is the *reactivation of reference* that we already mentioned in the prior section regarding mouthings (see 4). Such mouthings may be a noun or a name. Typically, a sign such as the noun TORTOISE can be used for the reactivation of a non-salient character. Another possibility is the use of fingerspelling. The simultaneous reactivation of reference by the narrator occurs most frequently when a matrix clause is missing and the signer has already shifted into the A-RS.

Adding supplementary information is the third function found in our data. This category comprises aspects of narration that are expressed mostly through classifiers, but also lexical signs may be of significance. We differentiate between three sub-categories: (i) *supplementary information about the same protagonist,* (ii) *supplementary information about a different protagonist*, and (iii) *general supplementary information* such as, for example, the sign HOT for the portrayal of the weather conditions. Of particular interest is the *addition of supplementary information about the same protagonist* in (9) taken from the fable *The lion and the mouse*.

(9) **Protagonist** *walk---|*
   rs
Narrator CLIMB-CL:4legs-over-the-face

The meaning conveyed by the sequence in Figure 8 can be paraphrased as 'the mouse is walking and climbing over the nose of the lion.' The first picture shows the A-RS$_{mouse}$ that represents the *walking*. Here, the facial expression of sharp lips is quite salient. The second picture illustrates an interesting layering of the classifier construction CLIMB-CL:4legs-over-the-face in combination with the facial expression of the mouse and the body or rather the nose of the lion. It is a case of partially mapping both the subject and object onto the signer's body. The facial expressions and the non-dominant hand belong to the mouse (subject) and the nose of the signer represents the nose of the lion (object) combined with the added information CLIMB-CL:4legs-over-the-face signed with the dominant hand by the narrator. The understanding of the overall meaning of the proposition that is intended by the signer requires these three sources of information. Another example for *adding supplementary information about the same protagonist* is given in (10).

| *walk* | CLIMB-CL |

Figure 8. Partial mapping of both the subject and object onto the signer's body in A-RS and the simultaneous addition of supplementary information by the narrator (*The lion and the mouse*)

(10)   **Protagonist** *run-------------------------|* ‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾‾ rs
        Narrator   APPROACH-CL:mass

In the A-RS as illustrated in the first picture in Figure 9, the body, the hands, and the face represent the action, in this case the running of a neighbour. The second picture presents an interesting layering of different perspectives. Again, the facial expression is used for the A-RS$_{neighbour}$. But at the same time, the narrator provides additional information by a classifier construction with two hands which means that a mass moves in the boy's direction. Still, the body of the signer represents the body of the boy. Hence, the narrator expresses two kinds of information: first, he points out the direction which the neighbours are running to and second, the narrator shows that the A-RS that is naturally articulated with only one body, namely the signers body, stands for the action of more than one person, i.e, many neighbours.

| *run* | APPROACH-CL |
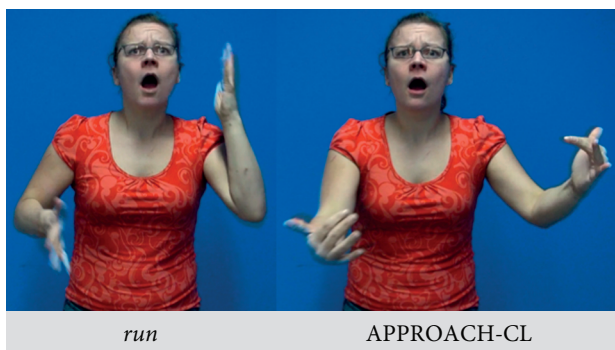
Figure 9. Partial mapping of both the subject and object onto the signer's body in A-RS and the simultaneous addition of supplementary information by the narrator (*The shepherd's boy and the wolf*)

Regarding the type of narrator's additions that we call *comment of attitude*, i.e. the adding of the narrator's opinion, we have not yet found an example in our elicited fables. However, there is an example provided by Metzger (1995: 262f.), where a signer is retelling a fight. Between two different perspective shifts using A-RS, the signer looks at the addressee and comments the situation by signing 'It was terrible.' This can also be analyzed in analogy to Clark and Gerrig's (1990: 768) *annotative aspects* of demonstrations in spoken languages. They explain that some "aspects of a demonstration are added as commentary on what is being demonstrated." This means the demonstrator might, for instance, smile or sneer during the demonstration as a comment on it. Within the category *comment of attitude,* the use of gestures and lexical signs may be used to add the opinion of the narrator.

The analysis of the four main functions of simultaneous additions of the narrator in A-RS has revealed that A-RS is a highly complex and multilayered phenomenon in sign languages. Fischer and Kollien (2010: 508) note that some types of A-RS are the most complex simultaneous constructions known in sign languages.

4. A unified account for role shift

4.1 Action role shift at the gesture-grammar interface

> We find dual messages in ASL provided by gesture and language. The existence of these dual messages strongly suggests a level of cognitive organization in language use higher than the grammatical level in which the two meanings are integrated. Thus, the simultaneous presence of gesture and language is not uniquely a spoken language phenomenon.   (Liddell & Metzger 1998: 694)

Fischer and Kollien (2010: 508) formulate as an open question whether A-RS is grammatical or gestural. We argue that it is a combination of both and that this means it should be analyzed as part of the language system in a unified account together with Q-RS along the continuum from purely gestural A-RS to purely grammatical Q-RS. Both extreme ends of the continuum are very rare in DGS and sign languages in general.

Considering the direct integration of classifiers in A-RS and the fact that A-RSs may function as introducing devices for Q-RSs, it becomes apparent that we deal with highly intertwined phenomena at the gesture-grammar interface. It is thus interesting to take into account the close interaction between Q-RS and A-RS. The latter can serve as a predicative structure introducing Q-RS. Looking at the numbers of a preliminary examination of two fables and three signers, it is obvious that all signers use A-RS to signal and introduce a following Q-RS (23% to 50%). Thus, some signers use up to half of their A-RSs for this sequential order (see Table 5).

Table 5. Percentage of A-RS that is followed by a Q-RS

Fable/Informant	A-RS total	A-RS & Q-RS sequential
1/A	20	6 = 30%
1/B	25	6 = 24%
1/C	17	4 = 24%
2/A	17	6 = 35%
2/B	21	7 = 33%
2/C	28	11 = 39%
3/A	12	6 = 50%
3/B	13	3 = 23%
3/C	30	8 = 27%
Total	183	57 = 31%

Furthermore, in our data we often observe A-RS as a sort of framing device for Q-RS. The same A-RS may occur preceding and following the Q-RS. These cases only involve simple and importantly identical A-RSs as in other doubling structures with certain predicates. Thus, it seems to be a regular doubling strategy for the framing of a Q-RS.

Arguing that A-RSs have predicative status was also suggested by Davidson (2015), who analyzed A-RS as predicative demonstrations. This would be in principle compatible with a semantic context shift operator. As we find complex combinations of sentential propositions, it is still hard to disentangle the predicative sequences. The difficulty of a clear-cut analysis arises, as role shift in itself allows for the shifted and non-shifted interpretation of specific temporal and local indexicals (cf. Quer 2005). In particular, the seemingly flexible nature of role shift in terms of the activation of different contexts and the shifting of indexicals respectively have lead to take alternative approaches based on spoken language phenomena such as Free Indirect Discourse (FID). These theories try to accout for the mixing of both of such contexts, the actual and the reported context. As mentioned above, for spoken languages and the establishment of a coherent context of narration, there are attempts to discuss the lexical semantics of certain shiftable elements (cf. Eckardt 2012) and based on a general shift together constraint (cf. Anand & Nevis 2004) apply these individually. Geurts and Maier (2005) tackle the problem from a different angle and propose an unquotation account based on conventionalized pragmatic constraints.

4.2   Integrating A-RS in an agreement analysis

The here presented agreement analysis of role shift in general is compiled to inclose the two main types Q-RS and A-RS. This is supported by Janzen (2004: 151) who outlines that

> the orientation of verbal elements, direction of movement paths, and location features with respect to perspective coding are critical in understanding how signers conceptualize events and the particular perspective they choose from several options to convey to the addressee.

Extending the analysis of Q-RS to A-RS, we need to disentangle the features that a sign/gesture agrees with in both types of roles shift. In Q-RS, nonmanuals agree with the loci of (the signer and) the addressee of the reported discourse, i.e, (3a and) 3b of context c. This is similar to agreement verbs in sign languages. In A-RS, nonmanuals similarly agree with referents of reported actions, usually using the loci of actor and addressee, i.e., 3a (and 3b) of context c, but not necessarily relying on them, as more prominently in A-RS, manual and nonmanual (facial) gestures may agree in the embodiment of the referent. One important point that needs to be discussed is whether subject/signer or object/addressee agreement is optional or obligatory like optional subject agreement in agreeing verbs, for instance. For pure Q-RS it seems that object locus agreement is obligatory as the break in eye gaze is such an important factor to signal quotation and can be used as an indicator of Q-RS without body and head movements. For pure A-RS, on the other hand, it seems that subject agreement is more important as it is sufficient to indicate the shift and enactment. Thus, pure A-RS is more like certain verbs that are lexically specified for subject agreement, like we find in subject/protagonist agreement in the way of 'body as subject' (cf. Meir et al. 2007). With a similar idea, Kegl (1986) analyzed this type of non-quotative role shift as a subject clitic, indicating the prominence of a role and agreeing with the respective person NP. For A-RS, body/facial expressions suffice for the agreement marking, especially if no concrete addressee is present as in fables when a protagonist acts out not straightforwardly directed towards an addressee. If a protagonist interacts with another character, loci are more frequently used also with A-RS.

Figure 10 sums up our classification of role shift. It is an extension and overall modification of an illustration by Hübl and Steinbach (2012). We illustrate that A-RS is also based on agreeing relations of the signers' body parts with particular discourse referents, thus semantic controllers. Instead of loci, the body parts themselves agree with individual features of characters, as this is possible with more gestural elements and facial expressions. Thus, Figure 10 is extended to capture not only Q-RS but also A-RS. An important difference between the two types of
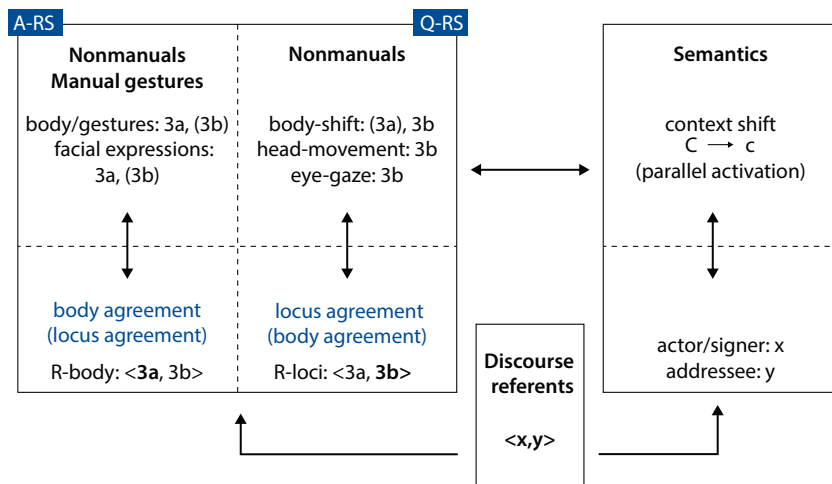
Figure 10. Extension and modifcation of the analysis of Q-RS by Hübl and Steinbach (2012)

role shift is the use of manual gestures and the leading strategy of 'body as subject' in A-RS. As illustrated, we find possible layerings of A-RS and Q-RS, thus, both forms may overlap and simultaneously appear. Hence, it seems obvious to handle role shift as a continuum between A-RS and Q-RS. As mentioned above, the continuum also changes from more gestural subject oriented marking like in A-RS to more grammatical object oriented marking in Q-RS. This mirrors the general process of grammaticalization from gesture to more functional elements, while, of course, both forms of role shift exist in parallel. We do not argue that role shift as a phenomenon is fundamentally different from comparable uses of gesture and eye gaze for enactment in non-signers, but that along the continuum from more gestural A-RS towards pure Q-RS, the features become grammaticalized and may be used solely grammatical to systematically mark quotations without gestural means. These features then agree with the subject locations and the object locations, which are the controllers (cf. Corbett 2006).

In addition, the extension to parallel focalization as agreement of different body parts with different referents, allows for the possibility of two simultaneously active contexts <C,c>. Both contexts can be referred to at the same time as different articulators (parts of the body such as torso, hands, and face) can indicate both protagonists and the narrator. The agreement relations can either be such that (i) parts of the signer demonstrate the action of the referent (manual and nonmanual), whereby the referents are represented by classifier handshapes restricted by language specific contraints, or (ii) comments relate to the narrator/signer (either signs or body and facial gestures). Thus, the context shift to c can be 'interrupted' and

the manual and nonmanual means can be used to shift back to C or simultaneously activate C. This is a general advantage of a shifting operator as proposed by Quer (2005) and it can also independently explain the shifting of certain indexicals (cf. Lillo-Martin 2012: 383). We assume the simultaneous activation of a third context to be possible (see example (3) above), but rare due to pragmatic constraints and the restrictions on the use of the signing space.

5.   Conclusion

The aims of this paper are twofold. First, we presented data from German Sign Language fables to illustrate the complex interplay of gestural and grammatical means for role shift. We focused on action role shift, which we called A-RS, and discussed parallel perspectivation and in which way two possible contexts may be activated simultaneously. We presented various categories of parallel perspectivation and how articulatory combinations of such simultaneous layerings unfold in A-RS. Specifically the various options of the narrator to interact with protagonists and the events such as what we call *reactivating referents* and *providing supplementary information and attitudes* have not been noted in detail yet. Thus, this in principle well-described phenomenon of multiple perspectives requires a meticulous classification to clearly disentangle various perspectives and their interaction.

Second, in an attempt to display the continuum that reaches from pure A-RS towards pure Q-RS and to unify the analysis within an agreement analysis framework, the features for A-RS may be triggered by agreement relations of semantic properties of agents such as the signer and the addressee of a shifted context. Thus, the continuum reaches from locus agreement for pure Q-RS to bodily referent agreement for A-RS. The clear linguistic restrictions on A-RS in terms of classifiers and the use of space and the systematic layering of different perspectives reveal the complex gesture-grammar-interface in DGS and the necessity to look at these phenomena from a formal perspective.

References

Aarons, D., Bahan, B., Kegl, J. & Neidle C. 1994. Subjects and agreement in American Sign Language. In *Perspectives on Sign Language Structure*. Papers from the Fifth International Symposium on Sign Language Research, Vol. 1, I. Ahlgren, B. Bergman & M. Brennan (eds), 13–27. Durham: International Sign Linguistics Association.

Aarons, D. & Morgan, R. Z. 2003. Classifier predicates and the creation of multiple perspectives in South African Sign Language. *Sign Language Studies* 3(2): 125–156.

Anand, P. & Nevins, A. 2004. Shifty Operators in Changing Contexts. *SALT XIV 20-37*, 20–37.

Aronoff, M., Meir, I., Padden, C. & Sandler, W. 2003. Classifier constructions and morphology in two sign languages. In *Perspectives on Classifier Constructions in Sign Languages*, K. Emmorey (ed.), 53–84. Hillsdale NJ: Lawrence Erlbaum Associates.

Benedicto, E. & Brentari, D. 2004. Where did all the arguments go? Argument-changing properties of classifiers in ASL. *Natural Language & Linguistic Theory* 22: 743–810.

Clark, H. H. & Gerrig, R. J. 1990. Quotations as demonstrations. *Language* 66: 764–805.

Corbett, G. G. 2006. *Agreement*. Cambridge: CUP.

Cormier, K., Smith S. & Zwets, M. 2011. *Show and tell: Framing and narrating constructed action in signed and spoken languages*. Paper presented at Complex Sentences and Beyond in Sign and Spoken Languages, Lichtenberg-Kolleg, Georg-August-Universitat Göttingen.

Cormier, K., Smith, S. & Zwets, M. 2013. Framing constructed action in British Sign Language narratives. *Journal of Pragmatics* 55: 119–139.

Cormier, K., Smith, S. & Sevcikova, Z. 2015. Rethinking constructed action. *Sign Language & Linguistics* 18(2): 167–204.

Crasborn, O., Mesch, J., Waters, D., Nonhebel, A., van der Kooij, E., Woll, B. & Bergmann, B. 2007. Sharing sign language data online. Experiences from the ECHO project. *International Journal of Corpus Linguistics* 12(4): 535–562.

Davidson, K. 2015. Quotation, demonstration, and iconicity. *Linguistics & Philosophy* 38(6): 477–520.

Dudis, P. G. 2004. Body partitioning and real-space blends. *Cognitive Linguistics* 15(2): 223–238.

Duncan, S. 2003. Gesture in language. Issues for sign language research. In *Perspectives on Classifiers constructions in Sign Languages*, K. Emmorey (ed.), 259–268. Mahwah NJ: Lawrence Erlbaum Associates.

Earis, H. & Cormier, K. 2013. Point of view in British Sign Language and spoken English narrative discourse: The example of 'The Tortoise and the Hare'. *Language and Cognition* 5(4): 313–343.

Eccarius, P. & Brentari, D. 2007. Symmetry and dominance: A cross-linguistic study of signs and classifier constructions. *Lingua* 117: 1169–1201.

Eckardt, R. 2012. Particles as speaker indexicals in free indirect discourse. In *Particles*, L. Hogeweg & E. McCready (eds), 99–119. Special issue of Sprache und Datenverarbeitung.

Emmorey, K. 2003. *Perspectives on Classifier Constructions in Sign Languages*. Mahwah NJ: Lawrence Erlbaum Associates.

Emmorey, K., Bellugi, U., Frederici A. & Horn, P. 1995. Effects of age of acquisition in grammatical sensitivity: Evidence from on-line and off-line tasks. *Applied Psycholinguistics* 16, 1–23.

Engberg-Pedersen, E. 1993. *Space in Danish Sign Language*. Hamburg: Signum.

Engberg-Pedersen, E. 1995. Point of view expressed through shifters. In *Language, Gesture, and Space*, K. Emmorey & J. Reilly (eds), 133–154. Hillsdale NJ: Lawrence Erlbaum Associates.

Fauconnier, G. & Turner, M. 1996. Blending as a central process of grammar. In *Conceptual Structure, Discourse and Language*, Goldberg, A. (ed.), 113–130, Stanford, CA: CSLI Publications.

Fischer, R. & Kollien, S. 2010. Gibt es Constructed Action in Deutscher Gebärdensprache und in Deutsch (in der Textsorte Bedeutungserklärung)? *Das Zeichen* 86: 502–510.

Geurts, B. & Maier E. 2005. Quotation in context. *Belgian Journal of Linguistics* 17(1): 109–128.

Graumann, C F. 2002. Explicit and implicit perspectivity. In *Perspective and Perspectivation in Discourse* [Human Cognitive Processing 9], C. F. Graumann & W. Kallmeyer (eds), 25–39. Amsterdam: John Benjamins.

Graumann, C. F. & Kallmeyer, W. 2002. Perspective and perspectivation in discourse. An introduction. In *Perspective and Perspectivation in Discourse* [Human Cognitive Processing 9], C. F. Graumann & W. Kallmeyer (eds), 1–11. Amsterdam: John Benjamins.

Grose, D., Wilbur, R., & Schalber, K. 2007. Events and telicity in classifier predicates: A reanalysis of body part classifier predicates in ASL. *Lingua* 117: 1258–1284.

Günthner, S. 1999. Polyphony and the 'layering of voices' in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics* 31: 685–708.

Günthner, S. 2002. Perspectivity in reported dialogues. The contextualization of evaluative stances in reconstructing speech. In *Perspective and Perspectivation in Discourse* [Human Cognitive Processing 9], C. F. Graumann & W. Kallmeyer (eds), 347–374. Amsterdam: John Benjamins.

Herrmann, A. & Steinbach, M. 2010. Eine neue Perspektive auf Role Shift in Deutscher Gebärdensprache (DGS). Perspektivwechsel als nichtmanuelles Kongruenzphänomen. *Das Zeichen* 84: 112–119.

Herrmann, A. & Steinbach, M. 2012. Quotation in sign languages. A visible context shift. In *Quotatives. Cross-linguistic and Cross-disciplinary Perspectives* [Converging Evidence in Language and Communication Research 15], I. Buchstaller & I. van Alphen (eds), 203–228. Amsterdam: John Benjamins.

Hübl, A. 2013. Role shift, indexicals and beyond – New evidence from German Sign Language. In *Proceedings from the 13th meeting of the Texas Linguistics Society*, L. C. Geer (ed.), 1–11.

Hübl, A. 2014. Context shift (im)possible – Indexicals in German Sign Language. In *Proceedings of ConSOLE XXI*, M. Kohlberger, B. Kate & E. Dutton (eds), 171–183. Leiden: Sole.

Hübl, A. & Steinbach, M. 2012. *Quotation across modalities. Shifting contexts in sign and spoken languages*. Paper presented at the workshop Quotation: Perspectives from Philosophy and Linguistics, Bochum.

Janzen, T. 2004. Space rotation, perspective shift, and verb morphology in ASL. *Cognitive Linguistics* 15(2): 149–174.

Kegl, J. 1986. Clitics in American Sign Language. In *Syntax and Semantics 19: The Syntax of Pronominal Clitics*, H. Borer (ed.), 285–309. New York: Academic Press.

Keyser R. & Larson-Hall, J. 2005. What does the critical period really mean? In *Handbook of Bilingualism. Psycholinguistic Approaches*, J. F. Kroll & A. M. B. de Groot (eds), 88–108. Oxford: OUP.

Leeson, L. & Saeed, J. I. 2007. Conceptual blending and the windowing of attention in simultaneous constructions in Irish Sign Language. In *Simultaneity in Signed Languages. Form and Function*, M. Vermeerbergen, L. Leeson & O. Crasborn (eds), 55–72. Amsterdam: John Benjamins.

Liddell, S. K. 1980. *American Sign Language Syntax*. The Hague: Mouton.

Liddell, S. K. 1995. Real, surrogate, and token space: Grammatical consequences in ASL. In *Language, Gesture, and Space*, K. Emmorey & J. Reilly (eds), 19–41. Hillsdale NJ: Lawrence Erlbaum Associates.

Liddell, S. K. 1998. Grounded blends, gestures, and conceptual shifts. *Cognitive Linguistics* 9(3): 283–314.

Liddell, S. K. 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: CUP.

Liddell, S. K. & Metzger, M. 1998. Gesture in sign language discourse. *Journal of Pragmatics* 30: 657–697.

Lillo-Martin, D. 1995. The point of view predicate in American Sign Language. In *Language, Gesture, and Space*, K. Emmorey & J. Reilly (eds), 155–170. Hillsdale NJ: Lawrence Erlbaum Associates.

Lillo-Martin, D. 2012. Utterance reports and constructed action. In *Sign Language. An International Handbook*, R. Pfau, M. Steinbach & B. Woll (eds), 365–387. Berlin: Mouton de Gruyter.

Lillo-Martin, D. & Klima, E.S. 1990. Pointing out differences: ASL pronouns in syntactic theory. In *Theoretical issues in sign language research*, S. D. Fischer & P. Siple (eds), 191–210. Chicago. University of Chicago Press.

Maier, E. 2015. Quotation and unquotation in free indirect discourse. *Mind & Language* 30(3): 345–373.

Meir, I., Padden, C. A., Aronoff, M. & Sandler, W. 2007. Body as subject. *Journal of Linguistics* 43: 531–563.

Metzger, M. 1995. Constructed dialogue and constructed action in American Sign Language. In *Sociolinguistics in Deaf Communities*, C. Lucas (ed.), 255–271. Washington DC: Gallaudet University Press.

Morford, J. P. 2004. Der Altersfaktor im Gebärdenspracherwerb: Eine Neuinterpretation. *Das Zeichen* 66: 85–88.

Morgan, G. 1999. Event packaging in British Sign Language discourse. In *Story Telling and Conversation. Discourse in Deaf Communities*, E. A. Winston (ed.), 27–58. Washington DC: Gallaudet University Press.

Morgan, G. & Woll, B. 2002. *Directions in Sign Language Acquisition* [Trends in Language Acquisition Research 2]. Amsterdam: John Benjamins.

Morgan, G. & Woll, B. 2003. The development of reference switching encoded through body classifiers in British Sign Language. In *Perspectives on Classifiers Constructions in Sign Languages*, K. Emmorey (ed.), 297–310. Mahwah NJ: Lawrence Erlbaum Associates.

Morgan, G. & Woll, B. 2007. Understanding sign language classifiers through a polycomponential approach. *Lingua* 117: 1159–1168.

Padden, C. 1986. Verbs and role-shifting in American Sign Language. In *Proceedings of the Fourth National Symposium on Sign Language Research and Teaching*, C. Padden (ed.), 44–57. Silver Spring MD: National Association of the Deaf.

Parrill, F. 2010. Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes* 25(5): 650–668.

Perniss, P. 2007a. Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua* 117: 1315–1338.

Perniss, P. 2007b. Locative functions of simultaneous perspective constructions in German Sign Language narratives. In *Simultaneity in Signed Languages. Form and Function* [Current Issues in Linguistic Theory 281], M. Vermeerbergen, L. Leeson & O. Crasborn (eds), 27–53. Amsterdam: John Benjamins.

Pfau, R. & Quer, J. 2010. Nonmanuals. Their grammatical and prosodic roles. In *Sign Languages*, D. Brentari (ed.), 381–402. Cambridge: CUP.

Pyers, J. E. & Senghas, A. 2007. Reported action in Nicaraguan and American Sign Languages. Emerging versus established systems. In *Visible Variation. Comparative Studies on Sign Language Structure*, P. Perniss, R. Pfau & M. Steinbach (eds), 279–302. Berlin: Mouton de Gruyter.

Quasthoff, U. M. 2002. Global and local aspects of perspectivity. In *Perspective and Perspectivation in Discourse* [Human Cognitive Processing 9], C. F. Graumann & W. Kallmeyer (eds), 323–346. Amsterdam: John Benjamins.

Quer, J. 2005. Context shift and indexical variables in sign languages. In *Proceedings of Semantics and Linguistics Theory 15*, E. Georgala & J. Howell (eds), 152–168. Ithaca NY: CLC Publications.

Quer, J. 2011. Reporting and quoting in signed discourse. In *Understanding Quotation*, E. Brendel, J. Meibauer & M. Steinbach (eds), 277–302. Berlin: Mouton de Gruyter.

Quinto-Pozos, D. 2007. Can constructed action be considered obligatory? *Lingua* 117(7): 1285–1314.

Quinto-Pozos, D. & Mehta, S. 2010. Register variation in mimetic gestural complements to signed language. *Journal of Pragmatics* 42: 557–584.

von Randow, S. 2016. Sprachlernerfahrungen erwachsener hörender L2-/M2-Lerner_innen bezüglich Constructed Action in Deutscher Gebärdensprache – Teil I: Ausgangslage und Methode. *Das Zeichen* 30(102): 116–127.

Schembri, A., Jones, C. & Burnham, D. 2005. Comparing action gestures and classifier verbs of motion: Evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners' gestures without speech. *Journal of Deaf Studies and Deaf Education* 10(3): 272–290.

Schick, B. 1990. Classifier predicates in American Sign Language. *International Journal of Sign Linguistics* 1: 15–40.

Schlenker, P. 2017. Super monsters I. Attitude and Action Role Shift in sign languages. *Semantics and Pragmatics*. 10(9).

Speas, M. & Tenny, C. 2003. Configurational properties of point of view roles. In *Asymmetry in Grammar* [Linguistik Aktuell/Linguistics Today 57], A. M. di Sciullo (ed.), 315–344. Amsterdam: John Benjamins.

Stec, K. 2012. Meaningful shifts. A review of viewpoint markers in co-speech gesture and sign language. *Gesture* 12(3): 327–360.

Stec, K., Huiskes, M. & Redeker, G. 2016. Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics* 104: 1–17.

Supalla, T. 1986. The classifier system in American Sign Language. In *Noun Classes and Categorization* [Typological Studies in Language 7], C. Craig (ed.), 181–214. Amsterdam: John Benjamins.

Sutton-Spence, R. 2007. Mouthings and simultaneity in British Sign Language. In *Simultaneity in Signed Languages. Form and Function* [Current Issues in Linguistic Theory 281], M. Vermeerbergen, L. Leeson & O. Crasborn (eds), 147–162. Amsterdam: John Benjamins.

Tannen, D. 2007[2]. *Talking Voices. Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: CUP.

Winston, E. A. 1991. Spatial referencing and cohesion in an American Sign Language text. *Sign Language Studies* 73: 397–410.

Zucchi, S. 2004. *Monsters in the visual mode?* Ms, Università degli Studi di Milano.

Zwitserlood, I. 2003. Classifiying Hand Configurations in Nederlandse Gebarentaal. PhD dissertation, University of Utrecht.

Zwitserlood, I. 2012. Classifiers. In *Sign Language. An International Handbook*, R. Pfau, M. Steinbach & B. Woll (eds), 158–185. Berlin: Mouton de Gruyter.

Index

In recent years, the focus of linguistic research has shifted from sentence to larger units such as text and discourse and accordingly from syntax to semantics and pragmatics. This has led to the development and application of corresponding discourse semantic and pragmatic theories such as, for instance, (S)DRT, Centering Theory, Accessibility Theory, QUD, Generalized Conversational Implicatures, Super Monsters and Gesture Semantics and new empirical approaches in the framework of experimental semantics and pragmatics or corpus linguistic discourse analysis. The contributions to this collected volume built on these developments and investigate the linguistic foundations of narration from various perspectives. The contributions address topics such as speech and thought representation, free indirect speech, information structure, anaphora resolution, co-speech gestures, classifier constructions as well as on role shift and constructed action. The volume provides new insights in the linguistic structures underlying narration in written, spoken, and sign languages from an experimental, developmental, historical, typological, and theoretical perspective. The contributions will appeal to theoretical linguists, sign language linguists, typologists, literary scholars, psycholinguists, and philosophers.

John Benjamins Publishing Company