

Demonstratives in discourse

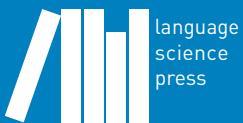
Edited by

Åshild Næss

Anna Margetts

Yvonne Treis

Topics at the Grammar-Discourse
Interface 6



Topics at the Grammar-Discourse Interface

Editors: Philippa Cook (University of Göttingen), Anke Holler (University of Göttingen), Cathrine Fabricius-Hansen (University of Oslo)

In this series:

1. Song, Sanghoun. Modeling information structure in a cross-linguistic perspective.
2. Müller, Sonja. Distribution und Interpretation von Modalpartikel-Kombinationen.
3. Bueno Holle, Juan José. Information structure in Isthmus Zapotec narrative and conversation.
4. Parikh, Prashant. Communication and content.
5. Balogh, Kata, Anja Latrouite & Robert D. Van Valin, Jr. (eds.) Nominal anchoring: Specificity, definiteness and article systems across languages.
6. Næss, Åshild, Anna Margetts & Yvonne Treis (eds.). Demonstratives in discourse.

Demonstratives in discourse

Edited by

Åshild Næss

Anna Margetts

Yvonne Treis



Næss, Åshild, Anna Margetts & Yvonne Treis (eds.). 2020. *Demonstratives in discourse* (Topics at the Grammar-Discourse Interface 6). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/282>

© 2020, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-286-0 (Digital)

978-3-96110-287-7 (Hardcover)

ISSN: 2567-3335

DOI: 10.5281/zenodo.4054814

Source code available from www.github.com/langsci/282

Collaborative reading: paperhive.org/documents/remote?type=langsci&id=282

Cover and concept of design: Ulrike Harbort

Typesetting: Yvonne Treis, Sebastian Nordhoff, Ahmet Bilal Özdemir

Proofreading: Alexia Fawcett, Amir Ghorbanpour, Andriana Koumbarou,

Barend Beekhuizen, Bev Erasmus, Christian Döhler, Carmen Jany, Franny

Vandervoort, Geoffrey Sampson, Jeroen van de Weijer, Joseph Lovestrond,

Kalen Chang, Lachlan Mackenzie, Madeline Myers, Natsuko Nakagawa, Ryan

Ka Yau Lai, Siva Kalyan, Sune Gregersen, Tom Bossuyt

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: X_EΛT_EX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin



Contents

Preface	iii
1 Introduction: Demonstratives in discourse Åshild Næss, Anna Margetts & Yvonne Treis	1
2 Beyond exophoric and endophoric uses: Additional discourse functions of demonstratives Ekkehard König	21
3 The use of manner demonstratives in discourse: A contrastive study of Wan (Mande) and Kambaata (Cushitic) Tatiana Nikitina & Yvonne Treis	43
4 Morphosyntactic and functional asymmetries in Vatlongos discourse demonstratives Eleanor Ridge	69
5 Tracking and recognitional use of Kalamang <i>opa</i> : Demonstrative of cognitive accessibility Eline Visser	103
6 <i>Stsíkiistsi ki stsíkiistsi</i> : The ubiquity of Blackfoot demonstratives in discourse Heather Bliss & Martina Wiltschko	123
7 Pilagá determiners and demonstratives: Discourse use and grammaticalisation Doris L. Payne & Alejandra Vidal	149
8 Referential shift potential of demonstrative pronouns – Evidence from text continuation Melanie Fuchs & Petra B. Schumacher	185

Contents

9 Psychologically distal demonstratives in Scandinavian are not “discourse new”	
Janne Bondi Johannessen	215
10 Space, contrast and joint attention: Demonstrative adverbs in Russian, Estonian and Finnish	
Tiina Nahkola, Maria Reile, Piia Taremaa & Renate Pajusalu	243
11 Manner deictics in quotative indexes of Finno-Ugric	
Denys Teptiuk	273
12 A typology of demonstrative clause linkers	
Holger Diessel & Merlijn Breunesse	305
Index	342

Preface

This volume arises out of a workshop on “Discourse Functions of Demonstratives” held at the University of Oslo on 14–15 June 2018. The workshop call for papers asked for abstracts on “functions of demonstratives and related forms as discourse-structuring devices beyond their already well-described reference tracking and discourse-deictic uses” and elicited a range of responses from which an interesting and stimulating two-day workshop was put together.

The editors would like to thank the Department of Linguistics and Scandinavian Studies at the University of Oslo for their generous financial support for organising the workshop, as well as the participants who contributed to an interesting discussion and convinced us that a volume of papers from the workshop would be a worthwhile endeavour. Many thanks to Jozina Vander Klok, who was a member of the workshop organising team in 2018. Special thanks to the anonymous reviewers of the chapters for their constructive criticism, to Melanie Burns for proof-reading the chapters, to Sebastian Nordhoff for his valuable support in formatting the book, and to the copy-editors who volunteer at Language Science Press for spotting remaining inconsistencies and typos in the manuscript.

During the final stages of preparation of this book, one of the contributors, Janne Bondi Johannessen, passed away after a cancer relapse. Janne was a colleague of Åshild Næss at the University of Oslo and shared with Yvonne Treis an interest in languages of Ethiopia. She was a dynamic and productive linguist with a keen interest in corpus research, who contributed greatly to linguistic research in a number of areas. She was director of the University of Oslo’s Text Laboratory from 1993 and coordinator of the NORHED-project “Linguistic capacity building in Ethiopia - Tools for the inclusive development of Ethiopia” (2014–2020). Janne was a great champion for accessibility and inclusiveness in academic work, and she supported many students in pursuing an academic career, both through her supervision and by involving them in projects at the Text Laboratory. She initiated linguistic capacity building and corpus building projects in different languages including Saami, Heritage Norwegian in the US and Ethiopian languages.

Janne completed the revisions of the paper included in this volume five months before she died. We dedicate this volume to her memory.

Chapter 1

Introduction: Demonstratives in discourse

Åshild Næss

University of Oslo

Anna Margetts

Monash University

Yvonne Treis

LLACAN (CNRS, INALCO)

1 Discourse markers and their functions

Over the last decades, there has been extensive discussion in the typological literature of the functions and uses of demonstratives. It is well established that demonstratives are not restricted to referring to items in situational use based on concrete spatial parameters, but that discourse deictic, anaphoric/tracking, and recognitional uses are also common, if not universal, functions of demonstratives (see Himmelmann 1996; 1997; and Diessel 1999 for systematic overviews). Studies have shown that many parameters beyond location and configuration of referents and speech-act participants play a role in demonstrative choice. In particular, directing the addressee's attention towards a target entity and prior knowledge of a referent either through the previous discourse or from the real world have been identified as relevant (see e.g. Burenhult 2003; Dawuda 2009; Diessel 2006; Enfield 2003; Hanks 1990; 1992; 2005; 2009; Küntay & Özyürek 2006; Özyürek 1998). The diachronic development from demonstratives to other types of markers with grammatical and discourse functions has also been extensively discussed (see again Himmelmann 1996; 1997; Diessel 1999).



This volume investigates discourse functions of demonstratives, that is, the type of functions demonstratives perform when they develop into discourse markers.¹ The notion of discourse marker is not very clearly defined, and the question thus arises which functions they comprise, how they can be described, and to what extent demonstratives and their functions match this description. In a broad-brush approach, discourse markers can be described as morphemes which deliver a meta-commentary on the discourse and establish and negotiate intersubjectivity. That is to say, discourse markers perform functions like directing attention as a means of establishing and maintaining joint attention, enabling the addressee to track participants through the discourse, indicating which parts of the discourse the speaker seeks to foreground and which are to be taken for granted, but also indicating the interlocutors' epistemic stance towards or evaluation of a particular portion of discourse (Englebretson 2007). Discourse markers communicate information like *hey, I'm starting something new here*, or *you know this already*, or *this is the important bit*, and *here is what I think about this*. They aid interlocutors in managing and navigating the discourse and in positioning themselves with regards to what is being said.

2 Origins of discourse markers: Lexicon vs. grammar

Discourse markers can originate from both lexical and grammatical elements, and as such they inhabit the interfaces between lexicon and discourse, but also between grammar and discourse. On the one hand, studies of discourse markers describe lexical items or phrases which are recruited into discourse functions, such as focus particles like *only*, *even*, *also* (König 1991), markers of discourse cohesion like *yeah-no* (Burridge & Florey 2002), or connectives like *and* (Schiffrin 2006) or *toujours* (Hansen 2006) to name but a few. On the other hand, discourse markers can be drawn from the domain of grammar and originate, for example, from person-based and spatial deictics (Margetts 2015). In some instances the distinction between lexicon vs. grammar as the source domains is blurred, as in the case of discourse markers like *y'know* and *I mean* (Schiffrin 1987), *I'm afraid* (Mazzon 2019), *voici/voilà* (Grenoble & Riley 1996), or Québec French *t'sais* 'you know' (Dostie 2009: 203), which originate from phrases composed of both lexical items and person or spatial deictics.

There are two ways in which discourse is seen as a distinct domain from grammar in much of the linguistic literature. The first is that discourse normally refers

¹A large number of terms are used in the literature for what we are here calling "discourse markers", e.g. "discourse particles", "pragmatic markers", "pragmatic particles"; see e.g. Degand & Simon-Vandenbergen (2011) and Heine (2013), and references therein.

to stretches of language longer than a sentence, while the traditional domain of grammar is the single sentence. The second is that discourse refers to utterances within a specific context, defined by such aspects as the speech-act participants and the relation between these, and the physical setting of the interaction; whereas grammar is traditionally considered as abstracted away from properties of any specific instance of use of a linguistic unit.

As a consequence, the terms “discourse marker” and “discourse functions” carry the connotation that the elements in question have functions outside and beyond their domain of origin, i.e. beyond their core function (be that lexical or grammatical), and in this sense the terms suggest functions “outside of grammar”. This functional distinction is mirrored in the formal properties often ascribed to discourse markers. For many researchers, they are defined by a certain degree of syntactic independence from their environment, i.e. they are not directly part of the grammatical structure or “not fully integrated into the syntactic structure of the utterance” (Dostie 2009: 209; see also e.g. Schiffrin 1987 and Heine 2013, among others).²

But the aspects of discourse vs. grammar outlined above also influence what researchers consider to be the primary functions of discourse markers. For example, Fraser (2006) considers discourse markers to be a subclass of the more general notion of pragmatic marker, and he restricts the term “discourse marker” to forms which signal “a relation between the discourse segment which hosts them and the prior discourse segment” (2006: 190), i.e. a purely structuring function in terms of relating parts of discourse to each other. Heine (2013), on the other hand, employs a broader definition which takes the principal function of discourse markers to be “to relate an utterance to the situation of discourse, more specifically to speaker-addressee interaction, speaker attitudes, and/or the organization of texts” (2013: 1211); that is, not only including the relations between parts of the discourse but also between the discourse and the speech act participants, and the knowledge treated as established or shared by them.

The understanding of grammar and discourse as largely distinct moreover presents a challenge for our investigation, as many of the discourse functions

²In Dostie’s terms, these forms have undergone a process of pragmatalisation (Dostie 2009: 203). She contrasts this with “non-discursive” uses of the same linguistic items (as in Québec French *t’sais* ‘you know’ as a discourse marker vs. as a verbal expression), thus underscoring the notion that discourse functions are distinct from “regular” functions of lexical and grammatical items. See also Degand & Simon-Vandenbergen (2011) and chapters in the special issue of *Linguistics* 49 for more discussion of whether the term grammaticalisation is applicable to the development of discourse markers, or whether a separate process of pragmatalisation should be assumed.

discussed in the contributions to this volume are covered by what has been described as “engagement” in the literature. This term was coined by Evans et al. (2018a: 1) to describe “grammaticalised means for encoding the relative mental directedness of speaker and addressee towards an entity or state of affairs”. This suggests that the type of functions which are typically considered discourse functions can, in individual languages, in fact be part of the grammar, and in this sense would not fall under the notion of discourse functions if these are defined as “outside of grammar”. In this particular case, it indicates that there are languages where grammar cannot be divorced from the need to situate the sentence in an utterance context which takes into account the relationship between the speaker, the addressee and their respective states of knowledge. Strictly speaking, this should come as no surprise to speakers of Western European languages which have grammaticalised definite articles, since definiteness is a tool for intersubjective coordination, a means used by the speaker to guide the addressee in identifying the correct referent, based on the speaker’s assumptions of the addressee’s state of knowledge and activation with respect to the referent (cf. Evans et al. 2018a: 117). It does, however, complicate the endeavour of defining a coherent set of discourse functions as separate from grammar.

The aim of this volume is to explore the kinds of functions that demonstratives can have in discourse cross-linguistically, and for this reason we take an inclusive approach to the notion of discourse marker, following authors such as Degand & Simon-Vandenbergen (2011), Heine (2013) and others. This means that we do not assume a strict division between grammar and discourse, and our use of the term “discourse function” includes markers of engagement as described by Evans et al. We employ the term in a broad functional sense to refer to forms which deliver a meta-commentary on the discourse and/or which establish intersubjectivity, independently of whether they are structurally fully integrated in the grammar of the language.

3 Demonstratives in grammar and discourse

The focus of this volume is on the discourse functions of demonstratives and the types of discourse markers which develop from such elements. The class of forms referred to as demonstratives is large and varied, both cross-linguistically and often also within individual languages. Traditional classifications of demonstratives (e.g. Diessel 1999; Dixon 2003) focus especially on the deictic distinctions made in demonstrative systems, and the morphology, syntax and grammaticalisation of demonstrative forms. Apart from exophoric functions, where

demonstratives identify entities in the physical context of the speech situation (and are typically accompanied by pointing gestures), it is well established that they regularly perform endophoric functions, where they refer to entities within the discourse. Such endophoric uses include introducing, identifying and tracking referents across the discourse (see e.g. Himmelmann 1996; Lichtenberk 1996; Dawuda 2009; Margetts 2020). Arguably, then, demonstratives are by their very nature situated at the grammar-discourse interface and predisposed to take on functions in this realm. This is reflected in the discussion by Diessel (2006), who suggests that demonstratives as a class fall outside the lexicon-grammar distinction. He argues that they “must be kept separate from both content words and grammatical markers” and “constitute a unique class of linguistic expressions serving one of the most fundamental functions in language: In their basic use, they serve to coordinate the interlocutors’ joint focus of attention” (Diessel 2006: 464). Focus of attention is where the exophoric and endophoric uses of demonstratives meet, in the sense that both functions involve directing the addressee in identifying a referent as the one the speaker has in mind. From this perspective, the question of whether discourse functions can be separated from grammar may be largely irrelevant for demonstratives, because by definition they incorporate aspects of both.

If demonstratives bridge the lexicon-grammar distinction and are predisposed to take on discourse functions, this raises the question of whether these properties have an impact on the types of markers they develop into, and whether they differ from discourse markers derived from other sources. One possible difference is that demonstratives which take on discourse functions more commonly remain constrained by their sentence-grammatical functions. This is in contrast to the approaches cited above which suggest that discourse markers are generally freed from the syntactic constraints of the source elements. Across the contributions to this volume, demonstratives are commonly described as retaining aspects of their grammatical constraints. Adnominal demonstratives as discourse markers commonly still occur in noun phrases, manner demonstratives retain the distributional properties associated with their deictic use, etc. Indeed this is one of the key points made by Nikitina & Treis (chapter 3): the distributional properties of the original items constrain the types of semantic-pragmatic extensions they may undergo. A similar point is made by Ridge (chapter 4) for the Oceanic language Vatlongos and by Teptiuk (chapter 11) for a number of Finno-Ugric languages, as well as from a more general typological perspective by Diessel & Breunesse (chapter 12). As we discuss in §4, this does not, however, mean that demonstratives *cannot* expand their syntactic scope as they take on discourse functions.

4 Overview of discourse functions discussed in this volume

The contributions to this volume touch on a wide range of different discourse-related functions. The functions addressed and the structure of the discussion in each chapter is determined by the features of the language and by the authors' focus, rather than by a framework dictated across the volume. In this section we present an overview of certain discourse functions which emerged as common threads across several chapters, however it constitutes by no means an exhaustive listing of functions addressed across the individual languages.

Degand & Simon-Vandenbergen (2011) suggest that discourse markers can be arranged on a scale spanning from strictly relational at one end, to non-relational at the other end. Relational discourse markers indicate relations between discourse segments, while non-relational forms are "primarily subjective and intersubjective markers, with little or no linking function" (Degand & Simon-Vandenbergen 2011: 289). Arguably, these two scalar poles relate to the two aspects or notions of discourse described in §1 (which distinguish it from "grammar" in the traditional sense): (a) discourse as comprising stretches of multiple sentences, and (b) discourse as referring to utterances in a context defined by speech-act participants and the relations between them.

Relational functions of discourse markers therefore can be thought to relate to discourse as comprising stretches of multiple sentences which need to be linked to or separated from each other in some fashion. By contrast, the non-relational, intersubjective function arises from discourse as referring to utterances in a specific context, which is defined by the speech-act participants and their relations and interactions. Below, we start from this distinction in order to discuss the main types of discourse functions discussed in this volume, and the relationships between them.

In §4.1 we summarise some of the functions relating to marking text structure and cohesion. §4.2 addresses discourse functions around engagement and intersubjectivity, and in §4.3 we consider emotional deixis and affect. In §4.4 we review the non-deictic members of demonstrative paradigms and the types of functions they express. However, it should be noted that in many cases discourse markers combine functions from more than one of these domains and the sections are a convenient means of structuring our discussion rather than clear categorical distinctions.

4.1 Text-structure and cohesion

In this section we address a number of functions which provide meta-information on the structure of the discourse. We subsume here, for example, tracking of referents throughout the discourse, but also indicating which stretches of discourse belong together or form separate thematic groupings. Most of these functions are located towards the relational end of Degand & Simon-Vandenbergen's scale, but they can also bridge between relational and non-relational functions. This is the case, for example, when markers perform text-structuring and intersubjective functions at the same time (such as highlighting stretches of discourse as forming a unit and at the same time evaluating them as important or unexpected).

Functions related to reference tracking and reference shift are strongly relational in that they indicate the relationship between a current referential expression and a preceding (or following) one. They indicate whether expressions have the same referent or whether the referent has changed. Such functions are the focus of the studies by Fuchs & Schumacher (chapter 8) and by Nahkola et al. (chapter 10). Fuchs & Schumacher study what they call the referential shift potential of two types of demonstrative pronouns in German, i.e. to what extent they are used to change the referential structure of a narrative by highlighting a previously less prominent referent in the ongoing discourse. Nahkola et al. look at how demonstrative adverbs are used in the description and comparison of locations in an experimental setting in Russian, Estonian and Finnish. They find that in all three languages, the proximal demonstrative adverbs occur mostly in first mentions of the referential chain, whereas the distal ones occur when the referent is already activated. However, in Finnish, activated referents are more commonly indicated with the addressee-centred demonstrative. Finnish uses demonstrative pronouns more often for first mentions and precise identification of referents, whereas Russian uses bare noun phrases more in such cases. Reference tracking is also described as a function of the Vatlongos demonstrative *ak* (Ridge, chapter 4), of Kalamang *opa* (Visser, chapter 5), and of several demonstrative forms in Pilagá (Payne & Vidal, chapter 7). Two of the chapters also address the preferences of demonstrative forms for anaphoric vs. cataphoric uses. In Vatlongos (Ridge), it is the proximal demonstrative which is used anaphorically, and in fact the other members of the three-term deictic paradigm do not show extensions into discourse functions at all. In the Permic language Udmurt, demonstratives in quotative indexes most frequently show an anaphoric function for the distal form and a cataphoric function for the proximal form (Teptiuk, chapter 11). This confirms suggestions in the literature (König, chapter 2; König 2012: 26–27) that distal (or: non-proximal) forms are more likely to have anaphoric functions

whereas proximal forms are more likely to be cataphoric, at least when referring to stretches of discourse as in English *I can't believe he said that* vs. *I'll tell you this: it's going to be tough*.

The use of demonstratives as discourse-connective devices is similarly relational in that they indicate an explicit link between the preceding and the following part of the discourse. Clause-connecting uses are the focus of the chapters by König (chapter 2), Diessel & Breunesse (chapter 12), and Teptiuk (chapter 11). König considers the use of demonstratives as connectives expressing rhetorical relations (English *so*, *hereby*). Diessel & Breunesse analyse eight types of clause linkers that are frequently derived from demonstratives cross-linguistically: relative pronouns, linking and nominalising articles, quotative markers, complementisers, conjunctive adverbs, adverbial subordinate conjunctions, correlatives and topic markers. Teptiuk looks at how manner demonstratives are used in quotative indexes, i.e. to introduce a direct or indirect quote, in computer-mediated communication in five Finno-Ugric languages: Finnish, Estonian, Hungarian, Udmurt, and Komi. These languages differ in the number of manner deictics and the contrasts made in the system, as well as in terms of which of the forms are found in quotative indexes. The patterns of use in quotative indexes are shown to depend on how the manner demonstratives are used in non-quotative contexts. A clause-connecting function is also found with the Kambaata manner demonstratives *hittigúta* and *híttikk* (Nikitina & Treis, chapter 3) and the Pilagá classifier *da?* and demonstrative root *m?e* (Payne & Vidal, chapter 7).

Also relating to discourse structuring properties in a broad sense are the functions around discourse repair reported for Vatlongos (Ridge, chapter 4), where the speaker-proximal demonstrative is used to mark a constituent which follows after a hesitation. In another repair-related function, the contrastive demonstrative is used to modify lexical placeholders (“that thing”) which are used instead of more specific formulations.

The use of demonstratives to foreground or background particular stretches of discourse (or participants) can still be considered relational in that it pertains to the relationship between different portions of discourse. This is particularly clear in the use of the speaker-proximal demonstrative in the so-called tail-head linkage construction in Vatlongos (Ridge, chapter 4), where a situation or referent that is introduced as new information at the end of one section is repeated as given information at the beginning of the next section of discourse. The repeated information thus serves as background to which further new information can be added. In Vatlongos, the speaker-proximal demonstrative is often used in the repeated sections, with the second, backgrounded mention or a referent. This

holds especially for referents that remain prominent in the subsequent discourse and in this way the demonstrative also carries evaluative meaning.

Backgrounding functions are also not strictly referential: They introduce a link between the discourse and the speech situation in that backgrounding again involves the speaker's subjective evaluation of which aspects of the discourse are to be emphasised or de-emphasised. This is particularly clear with functions explicitly described as evaluative, i.e. as indicating some form of commentary on the narrative by the speaker, or as marking the noteworthiness or expectedness of a particular stretch of discourse. Bliss & Wiltschko (chapter 6) discuss such a function for Blackfoot, where noteworthiness is indicated by proximate, as opposed to obviative, inflection on the demonstrative. Nikitina & Treis (chapter 3) also mention evaluation as a function of the extraposition construction with manner demonstratives in Wan. In Blackfoot, the distinction between speaker-proximal and addressee-proximal forms can indicate the expectedness or unexpectedness of the content of a proposition, which also relates to foregrounding and backgrounding functions in the sense that particular aspects of the discourse are being highlighted. The Blackfoot speaker-proximal form can mark a proposition as expected, predictable, or familiar to the addressee, while the addressee-proximal form is used when the content is unexpected, surprising, or new to the addressee. These can be considered as both relational and nonrelational functions. The demonstratives contribute to text coherence, but again also encode evaluation and stance in the sense that not just the proposition itself, but the speaker's assessment of it (as being predictable or surprising) is categorized. Similarly, in Pilagá, the 'unseen' demonstrative can be used as a marker of the speaker's uncertainty (Payne & Vidal, chapter 7). The classifiers which combine with the demonstrative roots into complex determiners can have modal uses in Pilagá, with the 'unseen, absent' classifier indicating the ignorance, desires or intentions of the speaker rather than a realised event. These functions also relate to the expressions of affect and emotional deixis which we address in §4.3.

The function of managing turn-taking, by indicating whether a particular utterance should be taken to open a new turn or to close an ongoing turn, is less linked to internal discourse coherence and more to managing the speech situation and the interaction between interlocutors. Such functions are discussed in this volume by König (chapter 2), who looks at the idiomatic use of demonstratives as discourse-structuring devices. These can be divided into those which are primarily used in introductory, initiating conversational moves (*English Now then!*, *Here you are!*) and those which are primarily used in responsive, conclusive conversational moves (*So there!*, *That's that*).

4.2 Engagement and intersubjectivity

A number of the chapters discuss demonstratives with a recognitional function, which mark a referent as being identifiable on the basis of shared knowledge between speaker and addressee (Himmelmann 1996). Typically, in a recognitional use, “the speaker is uncertain whether or not the kind of information he or she is giving is shared by the addressee or whether or not this information will be sufficient in allowing the addressee to identify the intended referent” (Himmelmann 1996: 230). Establishing that the Kalamang form *opa* is basically a recognitional demonstrative is the main focus of Visser’s contribution (chapter 5). Interestingly, in Kalamang, the recognitional use extends to exophoric situations (‘there’s that person whom we both know’), whereas previous studies (e.g. Diessel 1999) have considered the recognitional function as a subtype of endophoric demonstrative use.

Recognitional functions are also discussed for Vatlongos, where it is the proximal demonstrative that appears in this use (Ridge, chapter 4), for Pilagá, which has a dedicated recognitional demonstrative root (Payne & Vidal, chapter 7), and for Norwegian, where the third person singular pronouns *han/hun* as demonstratives can have a recognitional function (Johannessen, chapter 9).

The recognitional function typically pertains to the identification of a nominal referent, and so is linked to adnominal demonstratives. As discussed above, demonstratives commonly retain varying aspects of their grammatical constraints and functions when they develop into discourse markers and so adnominal demonstratives typically remain adnominal when their usage is extended, and they take on functions scoping over the noun phrase in which they appear. However, changes in scope are also attested, and scope extensions from single words or phrases to entire clauses are commonly described (see e.g. the contributions to Bril 2010 and Margetts 2016; 2019; 2020). An interesting case of scope change is discussed by Bliss & Wiltschko for the Blackfoot “untranslatable demonstratives” (chapter 6). These forms have undergone a shift in scope in the direction towards greater syntactic independence – which is considered characteristic of discourse markers by many researchers, as discussed in §1: They no longer take a nominal complement and are frequently clause-initial. In parallel to this structural change, functionally they scope over the whole proposition, marking distinctions of epistemicity, such as whether the event described is considered to be expected or surprising in the context of the unfolding discourse. Bliss & Wiltschko describe these functions as recognitional use, in parallel to the use of adnominal demonstratives as marking that the intended referent has to be identified via specific but presumably shared knowledge (Himmelmann 1996: 207). However, the Blackfoot

forms in this particular usage differ from the canonical recognitional function in that they do not pertain to the identification of a referent, but to the degree of predictability of an event. If we make explicit the distinction between demonstratives that scope over referents and those that scope over propositions, we can thus understand the function of the untranslatable Blackfoot demonstratives as being similar, but not identical, to the canonical recognitional use. While the latter pertains to knowledge about a referent, the Blackfoot demonstratives function to manage intersubjective knowledge about events in the discourse. In their discussion of engagement, Evans et al. (2018a: 134) note that “[t]here are some important differences between engagement as it can apply to objects (especially objects that are present in the speech situation) and as it applies to events and situations”. The Blackfoot data shows that demonstratives are not only able to encode object-related engagement, but also the event and situation-related type.³

4.3 Emotional deixis and affect

The idea that demonstratives can indicate “emotional deixis”, invoking a sense of solidarity and shared emotional involvement, goes back to Lakoff (1974). We are not aware of any systematic cross-linguistic studies of this use of demonstratives, although Šimík (2016) cites work on Czech, German and Japanese as well as English. Potts & Schwarz (2010) link the emotional deictic function of demonstratives to their spatial-deictic function, and suggest that “affective uses of *this* arise from its more basic meaning as a marker of spatio-temporal proximity ... If this is correct, then we expect affectivity for proximal demonstratives quite generally across languages” (Potts & Schwarz 2010: 25).

The extension of spatial deictics to affective meaning and to coding referents one feels positively about as ‘close’, and those one feels negatively about as ‘far’ is attested in a range of language (see e.g. Farr & Whitehead 1981 and Pryor 1990: 19 on the Papuan languages Korafe and Botin). Interestingly, however, there are three languages in this volume with demonstrative uses which could be described as emotional deixis – Pilagá, Vatlongos, and Norwegian – and in two of them the relevant forms do not in fact encode a deictic value.⁴ This perhaps suggests that the emotional use of demonstratives is not exclusively linked to deixis and

³This possibility is not explicitly discussed by Evans et al., though they do note that one item in the evidential paradigm of the Papuan language Foe might be related to a distal demonstrative (Evans et al. 2018b: 11).

⁴The Blackfoot “untranslatable” demonstratives can also have an emotive function, but this is achieved through lengthening of the vowel of the demonstrative rather than by the demonstrative in and of itself (Bliss & Wiltschko, chapter 6: §4.4).

distance but also to the inherently intersubjective function of demonstratives as seeking to establish a common ground between speaker and addressee, as suggested by Acton & Potts (2014).

Of the three mentioned languages, only in Pilagá, the ‘near’ vs. ‘far’ deictic classifiers encode emotivity in a way that can be linked to their spatial-deictic function: Characters with which the narrator sympathises take the ‘near’ form, whereas those who are not deserving of sympathy take the ‘far’ form (Payne & Vidal, chapter 7: §3). In Vatlongos, however, it is the non-deictic contrastive demonstratives in *-e* which take on the function of negative affect, being linked to enemies and troublemakers and to times of fear and distress (Ridge, chapter 5: §4.3).

In Norwegian, the forms analysed by Johannessen (chapter 9) as “psychologically distal” demonstratives are in fact third person pronouns used as adnominal demonstratives, and, again, they do not in themselves encode deictic values (although they can be combined with demonstrative adverbs, as in *han fyren der* ‘that guy there’). As Johannessen shows, in constructions without the adverb, i.e. where no deictic contrast is being made, the pronoun typically indicates the speaker’s negative attitude to the referent.

It is worth noting that the Norwegian construction in question also has recognitional uses, i.e. it is used to draw attention to shared knowledge between the speaker and addressee. In Vatlongos, on the other hand, the contrastive demonstratives in *-e* do not occur with a recognitional function (though the proximal demonstrative in *=ak* does), and so its extension to marking negative affect must be explained by some other means – presumably as being rooted in the contrastive function itself. Ridge states that the Vatlongos contrastive demonstrative is “closely linked to the breaking of social conventions, especially taboos and concepts of respect and obedience”, and perhaps this arises from the notion of contrast as indicating that something is different or diverges from norms or expectations. It is clear that there is considerable potential for further exploration of this aspect of demonstrative use from a cross-linguistic perspective.

4.4 Non-deictic demonstratives

Demonstrative systems are generally defined and described in terms of distance distinctions, either degrees of distance from the speaker or other deictic centre (so-called distance-based systems) or proximity vs. distance to the speaker and/or the addressee in the speech situation (person-based systems). However, demonstrative systems quite commonly include forms which do not encode deictic contrast, but which can nevertheless be analysed as demonstratives, because

they are in a paradigm with forms that do distinguish deictic contrasts. We are not primarily referring here to demonstratives which are spatially neutral or unmarked, as discussed by Levinson (2018), nor to demonstratives which *in addition to* deictic oppositions also express further functions. Our focus here is on items which have a specialised meaning and which does not involve spatial reference. Commonly such non-deictic members of demonstrative paradigms are dedicated to functions like anaphora or contrast, which are commonly attested as additional functions of deictic demonstratives cross-linguistically. An example is the well-documented Latin demonstrative paradigm, which includes a dedicated anaphoric form. We are not aware of any systematic cross-linguistic studies of such non-deictic members of demonstrative paradigm but the languages represented in this volume suggest that they are fairly common.

Systems contrasting a proximal, a distal and an ‘unseen’ demonstrative, as described for Pilagá (Payne & Vidal, chapter 7), are relatively common in Native American languages (Diessel 1999: 41–42). As far as dedicated recognitional demonstratives are concerned, they are attested at least in a number of Australian languages (Himmelmann 1996: 231–233). In the present volume such a form is attested in Kalamang, a Papuan (West Bomberai) language, as mentioned in §4.2.

Ridge (chapter 4) describes a dedicated contrastive demonstrative in Vatlongos, an Oceanic language. Contrastive uses of demonstratives are also discussed by Meira & Terrill (2005). Treis (2020: 351) suggests that the presence of a dedicated contrastive demonstrative in a language is cross-linguistically rare, but cites examples from the Omotic language Gamo and the Cushitic language Alaaba, as well as the latter’s close relative Kambaata, which is also discussed in the present volume.⁵

Besides demonstratives dedicated to non-deictic functions, there are other types of forms and paradigms which can make it difficult to decide how the class of demonstratives should be delimited and defined in a given language, let alone cross-linguistically. In some languages spatial deictics interact with different form classes that encode other types of functions. In this volume, this is the case particularly in Pilagá (Payne & Vidal, chapter 7), where demonstrative roots with a core deictic meaning interact with classifiers in complex ways. In many cases, this involves the encoding of posture and spatial orientation of referents (e.g. as upright, vertical, non-extended), a function not generally associated with demonstrative systems.

As mentioned in §4.1, in Blackfoot, demonstratives can take the proximate/obviative inflection also found on nouns (Bliss & Wiltschko, chapter 6), and this

⁵The Nama demonstrative *náú*, described in Anderson & Keenan (1985: 286) as distal yet “only used for contrastive purposes”, seems another likely candidate.

distinction has functions related to salience and noteworthiness of the narrated event. It is the combination of the demonstrative form and proximate/obviative distinction that performs this discourse function, and one might thus debate to what extent this function can be attributed to demonstratives as such.

5 Overview of chapters

The chapters in this volume cover a broad range of languages from across the world, and a range of source constructions and discourse functions. The demonstratives described here belong to systems of very different degrees of complexity (see the morphologically highly complex systems in the Indigenous languages of the Americas, as described by Bliss & Wiltschko and Payne & Vidal), to different morphosyntactic types (see e.g. Diessel & Breunesse) and ontological types (see the discourse functions of manner demonstratives in König, Teptiuk, and Nikitina & Treis). In many languages it is observed that demonstratives of different deictic specifications take diverging paths and develop clearly distinct discourse functions (see König, Ridge, Nahkola et al., Teptiuk). Several of the studies also point out that the syntactic restrictions on individual forms contribute to shaping their extension into the domain of discourse.

KÖNIG looks at how one can analyse discourse functions of demonstratives which go beyond the established distinction between endophoric and exophoric uses, by investigating which aspect of the core meaning of the demonstrative is lost when its usage is extended: the deictic component, the ontological component, or both. The chapter focuses on demonstratives of manner, quality and degree, which have received relatively little attention in the literature.

NIKITINA & TREIS examine how manner demonstratives are used in discourse in the African languages Kambaata (Cushitic) and Wan (Mande). They show that, while the core uses of manner demonstratives in these two unrelated languages are strikingly similar, their extended uses differ significantly. The authors attribute these differences to the fact that manner demonstratives belong to different syntactic categories in the two languages.

RIDGE discusses the demonstrative system in the Oceanic language Vatlongos. In the spatial domain, Vatlongos makes a three-way distinction between speaker-proximal, addressee-proximal, and distal demonstratives. However, this three-way distinction does not carry over to the use of demonstratives in discourse; here, the language shows a two-way distinction between a speaker-proximal and a contrastive form. Ridge surveys the discourse uses of these two forms, and proposes that the asymmetry between the broad range of discourse functions

found with the proximal clitic *ak* and the restricted uses of contrastive *-e* can be partly accounted for by the differences in their syntactic distribution.

VISSEER describes the functions of the demonstrative *opa* in the West Bomberai language Kalamang. This form does not encode any deictic distinctions and can only be used adnominally. Visser considers *opa* to indicate the cognitive accessibility of a referent, either via previous mention or in a recognitional sense, by indicating that the addressee has engaged with the referent in real life outside the discourse. The form *opa* is homophonous with a temporal adverb meaning ‘just now, earlier’, a pattern which is also found in some other languages of Indonesia. Visser suggests a possible grammaticalisation path from adverb to demonstrative, based on the fact that recent events are accessible to the minds of those who have witnessed them.

BLISS & WILTSCHKO examine a set of forms which they label “untranslatable demonstratives” in Blackfoot, which involve no clear nominal referent and do not make any truth-conditional contribution to the content of the utterance. They show that the untranslatable demonstratives are involved in a number of discourse functions, such as marking a proposition as expected or familiar versus unexpected or new, marking events as salient and noteworthy, and, in combination with vowel lengthening, indicating the speaker’s emotive attitude or personal reaction towards the content of the utterance.

PAYNE & VIDAL describe the demonstrative system in the Guaykuruan language Pilagá, which combines demonstrative roots with classifiers and gender markers in a complex system of determiners. The classifiers and demonstratives can have temporal or modal as well as spatial interpretations. Moreover, the ‘vertically extended’ classifier has been grammaticalised into a clausal subordinator, whereas the demonstrative *m?e*, which can contrast with the proximal and distal forms to indicate medial distance, but which frequently seems to be distance-neutral, also functions as a relativiser. This seems to be linked to an association of *m?e* with information assumed by the speaker to be already activated for the addressee, in contrast to the recognitional demonstrative *naqae*, which functions as an instruction to activate identifiable information.

The study by FUCHS & SCHUMACHER looks at the use of two types of demonstratives in German – *der/die/das* vs. *dieser/diese/dieses* – and compares these to each other and to personal pronouns in terms of the referents they typically refer to anaphorically. The results of an experimental task show that the referential shift potential of different forms appears to be modulated by the interpretive preference of the pronouns. It is not the case that demonstrative pronouns always initiate a referential shift, while personal pronouns maintain the previ-

ously established prominence ranking. Rather, when the personal pronoun is interpreted as referring to the less prominent character, it can initiate a referential shift in favour of this less prominent character. The results moreover show different referential dynamics for these forms: Demonstrative *dieser* appears to shift prominence to a referent only briefly, whereas demonstrative *der* seems to be associated with continuing high prominence of a previously less prominent referent.

JOHANNESSEN looks at the use of third person pronouns as demonstratives in Norwegian in expressions like *han læreren* (3SG.M teacher.DEF.M) ‘that teacher’. These have been argued to have a recognitional use, introducing referents as “discourse new and hearer old”. Johannessen, however, shows that such an analysis cannot account for the majority of cases in her data, as such forms regularly occur in noun phrases with discourse-active referents. Instead, she argues that they are used to indicate psychological distance, defined as either a lack of familiarity with the referent, or a negative attitude towards the referent.

NAHKOLA, REILE, TAREMAA & PAJUSALU compare the use of demonstrative adverbs in Russian, Estonian and Finnish in the description and comparison of locations in an experimental setting. They find parallels in the functions of proximal vs. distal forms, but also differences in terms of which formal class of morphemes (demonstrative pronouns, demonstrative adverbs, personal pronouns, etc.) is used for which function. The authors suggest that the differences between the languages may at least to some extent go back to the differences in paradigmatic structure and the syntactic potential of the forms available in each language.

TEPTIUK looks at how manner deictics are used in quotative indexes in computer-mediated communication in five Finno-Ugric languages: Finnish, Estonian, Hungarian, Udmurt and Komi. The languages differ in the number of manner deictics and in the semantic contrasts made in the system, as well as in which of the forms are found in quotative indexes. Teptiuk shows that the patterns of use of manner deictics as quotative indexes depend on their uses outside the quotative domain, and suggests that parallels can be found in languages outside the Finno-Ugric family.

Finally, DIESSEL & BREUNESSE look at how demonstratives grammaticalise into clause linkers cross-linguistically. Based on a sample of 100 languages, they look in detail at the types of constructions involving demonstratives from which each type of linker can be shown (or be hypothesized) to have grammaticalized. They emphasise the fact that the grammaticalisation of clause linkers is influenced by the syntactic properties of demonstratives in particular constructions. Moreover,

the authors make the point that not all demonstrative clause linkers are immediately derived from anaphoric demonstratives and discourse deictics: The various types of demonstrative clause linkers are historically related to each other, and these relationships are crucial for understanding the occurrence of demonstratives in certain clause-linkage constructions.

References

- Acton, Eric K. & Christopher Potts. 2014. That straight talk: Sarah Palin and the sociolinguistics of demonstratives. *Journal of Sociolinguistics* 18(21). 3–31.
- Anderson, Stephen R. & Edward L. Keenan. 1985. Deixis. In Timothy Shopen (ed.), *Language typology and grammatical description*, vol. 3: *Grammatical categories and the lexicon*, 259–308. Cambridge: Cambridge University Press.
- Bril, Isabelle (ed.). 2010. *Clause linking and clause hierarchy: Syntax and pragmatics*. Amsterdam: Benjamins.
- Burenhult, Niclas. 2003. Attention, accessibility, and the addressee: The case of the Jahai demonstrative *ton*. *Pragmatics* 13(3). 363–379.
- Burridge, Kate & Margaret Florey. 2002. *Yeah-no he's a good kid*: A discourse analysis of *yeah-no* in Australian English. *Australian Journal of Linguistics* 22(2). 149–171.
- Dawuda, Carmen. 2009. *Discourse functions of demonstratives and place adverbs with exophoric reference in Logea, an Oceanic language of Papua New Guinea*. Melbourne: Monash University. (PhD dissertation).
- Degand, Liesbeth & Anne-Marie Simon-Vandenbergen. 2011. Introduction: Grammaticalization and (inter)subjectification of discourse markers. *Linguistics* 49(2). 287–294.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1). 61–112.
- Dostie, Gaétane. 2009. Discourse markers and regional variation in French: A lexico-semantic approach. In Kate Beeching, Nigel Armstrong & Françoise Gadet (eds.), *Sociolinguistic variation in contemporary French*, 201–214. Amsterdam: Benjamins.
- Enfield, Nick. 2003. Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis. *Language* 79(1). 82–117.

- Englebretson, Robert. 2007. Stancetaking in discourse: An introduction. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 1–25. Amsterdam: Benjamins.
- Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018a. The grammar of engagement I: Framework and initial exemplification. *Language and Cognition* 10(1). 110–140.
- Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018b. The grammar of engagement II: Typology and diachrony. *Language and Cognition* 10(1). 141–170.
- Farr, Cynthia J. & Carl R. Whitehead. 1981. This, that, and the other: A study of Korafe demonstratives. *Language and Linguistics in Melanesia* 13(1–2). 64–80.
- Fraser, Bruce. 2006. Towards a theory of discourse markers. In Kerstin Fischer (ed.), *Approaches to discourse particles*, 189–204. Oxford: Elsevier.
- Grenoble, Lenore A. & Matthew Riley. 1996. The role of deictics in discourse coherence: French *voici/voilà* and Russian *vot/von*. *Journal of Pragmatics* 25. 819–838.
- Hanks, William F. 1990. *Referential practice: Language and lived space among the Maya*. Chicago, IL: University of Chicago Press.
- Hanks, William F. 1992. The indexical ground of deictic reference. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 43–76. Cambridge: Cambridge University Press.
- Hanks, William F. 2005. Explorations in the deictic field. *Current Anthropology* 46(2). 191–220.
- Hanks, William F. 2009. Fieldwork on deixis. *Journal of Pragmatics* 41(1). 10–24.
- Hansen, Maj-Britt Mosegaard. 2006. A dynamic polysemy approach to the lexical semantics of discourse markers (with an exemplary analysis of French *toujours*). In Kerstin Fischer (ed.), *Approaches to discourse particles*, 21–42. Oxford: Elsevier.
- Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6). 1205–1247.
- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.
- König, Ekkehard. 1991. *The meaning of focus particles: A comparative perspective*. London: Routledge.

- König, Ekkehard. 2012. Le rôle des déictiques de manière dans le cadre d'une typologie de la déixis. *Bulletin de la Société de Linguistique de Paris* 107(1). 11–42.
- Küntay, Aylin C. & Aslı Özyürek. 2006. Learning to use demonstratives in conversation: What do language specific strategies in Turkish reveal? *Journal of Child Language* 33(2). 303–320.
- Lakoff, Robin. 1974. Remarks on *this* and *that*. In Michael W. LaGaly, Robert A. Fox & Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, 345–356. Chicago, IL: Chicago Linguistic Society.
- Levinson, Stephen C. 2018. Introduction: Demonstratives: Patterns in diversity. In Stephen Levinson, Sarah Cutfield, Michael J. Dunn, N. J. Enfield & Sérgio Meira (eds.), *Demonstratives in cross-linguistic perspective*, 1–42. Cambridge: Cambridge University Press.
- Lichtenberk, Frantisek. 1996. Patterns of anaphora in To'aba'ita narrative discourse. In Barbara A. Fox (ed.), *Studies in anaphora*, 379–412. Amsterdam: Benjamins.
- Margetts, Anna. 2015. Person shift at narrative peak. *Language* 91(4). 755–805.
- Margetts, Anna. 2016. Addressee-based demonstratives as topic markers. Paper presented at the Annual Conference of the Australian Linguistics Society, Monash University, Melbourne, December 7–9, 2016.
- Margetts, Anna. 2019. Demonstratives in topic-related functions. Paper presented at the 11th Conference of Oceanic Linguistics, Nouméa, New Caledonia, October 7–11, 2019. <https://www.youtube.com/watch?v=mp3SOhZ-wVk>.
- Margetts, Anna. 2020. Different sources, same path: The development of addressee based deictics to markers of discourse status. In Keith Allan (ed.), *Dynamics of language changes: Looking within and across languages*, 1–19. Singapore: Springer Nature.
- Mazzon, Gabriella. 2019. Variation in the expression of stance across varieties of English. *World Englishes* 38(4). 593–605.
- Meira, Sérgio & Angela Terrill. 2005. Contrasting contrastive demonstratives in Tiriyó and Lavukaleve. *Linguistics* 43(6). 1131–1152.
- Özyürek, Aslı. 1998. An analysis of the basic meaning of Turkish demonstratives in face-to-face conversational interaction. In Serge Santi, Isabelle Guaïtella, Christian Cavé & Gabrielle Konopczynski (eds.), *Oralité et gestualité : Communication multimodale, interaction*, 609–614. Paris: Harmattan.
- Potts, Christopher & Florian Schwarz. 2010. Affective *this*. *Linguistic Issues in Language Technology* 3(5). 1–30.

- Pryor, John. 1990. Deixis and participant tracking in Botin. *Language and Linguistics in Melanesia* 21(1–2). 1–29.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Schiffrin, Deborah. 2006. Discourse marker research and theory: Revisiting *and*. In Kerstin Fischer (ed.), *Approaches to discourse particles*, 315–338. Oxford: Elsevier.
- Šimík, Radek. 2016. On pragmatic demonstratives: The case of pragmatic discourse anaphora in Czech. In Nadine Bade, Polina Berezovskaya & Anthea Schöller (eds.), *Proceedings of Sinn und Bedeutung*, vol. 20, 640–657.
- Treis, Yvonne. 2020. Presentatives in Kambaata from a Cushitic perspective. In Isabelle Leblie & Lameen Souag (eds.), *Du terrain à la théorie : Les 40 ans du LACITO. Actes du colloque international, 15–17 novembre 2016, Villejuif (France)*, 343–369. Villejuif: LACITO Publications.

Chapter 2

Beyond exophoric and endophoric uses: Additional discourse functions of demonstratives

Ekkehard König

Free University of Berlin

It is a well-known fact that demonstratives develop endophoric uses, i.e. anaphoric and cataphoric ones, as extensions of their basic exophoric (gestural) use. In these endophoric uses the relevant expressions relate retrospectively either to a part of the preceding, or prospectively to a part of the following discourse. These three use types and their role in the development of grammatical markers have been analyzed in a wide variety of studies for a large number of languages. Far less attention has been given so far to the analysis and systematization of additional discourse functions of demonstratives, in which they have either lost the deictic or the ontological component of their meaning, or even both. The present chapter analyzes such uses for a small subset of European languages, drawing a distinction between four types: (i) coordination of contrasting terms and their use to express quantification and vagueness, (ii) idiomatic combinations of basic demonstratives and their use as discourse-structuring devices, (iii) the use of demonstratives relating to evaluative rather than standard points of orientation and (iv) the basically anaphoric use of demonstratives as adverbial connectives enriched in meaning by rhetorical relations. The major focus of the study are atypical demonstratives (e.g. English *so*, *such*), expressing the ontological components of manner, quality and degree.

1 Introduction

It is generally acknowledged that in addition to their exophoric (deictic, gestural) use, demonstratives typically also have endophoric uses, which can be regarded as a first step in their further grammaticalisation. Building on some ty-



pological surveys (Diessel 1999; Dixon 2003) and on a variety of previous studies (König 2015; 2017; König & Nishina 2015), the present chapter aims to identify and analyse additional discourse functions of demonstratives that have been so far neglected in cross-linguistic studies. Since the identification of relevant discourse functions requires a high degree of familiarity with the relevant languages, the data for this study will be mainly taken from four European languages (English, German, French, Italian), as representatives of languages with few deictic differentiations. Data will also partly come from Japanese, as a representative of languages with a rich system of relevant differentiations. Moreover, a subset of largely neglected and “atypical” demonstratives denoting the ontological dimensions of manner, quality and degree (MQD-demonstratives: German *so*, *solch*; English *so*, *such*; French *ainsi*, *tel/pareil*, *tellement*) will receive specific attention. Four major types of uses that cannot simply be subsumed under the exophoric or endophoric (i.e. anaphoric and cataphoric) use types will be distinguished; these types of uses either lost these basic deictic and/or ontological meaning or enriched their meaning in such a way that a basic anaphoric function is only marginally visible in their new use as adverbial connectives. These uses are found in (a) coordination of contrasting terms in a demonstrative paradigm, (b) idiomatic combinations of basic demonstratives, (c) a transfer of the *origo* (i.e. the point of orientation) from the coordinates of the speech situation to some evaluative points, and (d) adverbial connectives derived from demonstratives. In addition to being identifiable on the basis of such formal properties, the relevant uses can also be characterised through the major changes leading to such uses: the loss of the deictic component of demonstratives leads to uses (a) and (c), whereas the loss of their ontological component is a salient feature of use (b). In (d) it is not the loss of a component but rather the contextual enrichment of the ontological component that underlies this use.

2 Exophoric and endophoric uses

Evidence from language learning (early acquisition), from language change (referent identified by gesture > referent identified in the preceding or following discourse), as well as from evolutionary hypotheses (“from gesture to grammar”, cf. Arbib 2012) have led to more or less general agreement that the deictic (gestural, exophoric) use of demonstratives is basic and that endophoric (anaphoric and cataphoric) uses are derived from this basic source.¹ The examples of adver-

¹This is as good a place as any to point out that the question of which phenomena are to be included under the term “deixis” is a matter of some debate, ranging from very restrictive approaches (cf. Kibrik 2011: 504ff.) to more encompassing ones (e.g. Gerner 2009: 70).

bial and adnominal demonstratives with a basically spatial meaning in (1) and (2) illustrate these uses.

- (1) a. *The restaurant over **there*** (+ pointing gesture) *is where we want to go.* (exophoric)
- b. *John has moved to Jakarta. Myself, I would not want to live **there**.* (anaphoric)
- c. ***Here** is what he said: “...”* (cataphoric)

- (2) a. ***That*** (+ pointing gesture) *book is exactly what I want.* (exophoric)
- b. *He offered me some advice, but I did not want **that**.* (anaphoric)
- c. *Let me tell you **this**:* “...” (cataphoric)

These examples show that it is basically, though not exclusively, the distal term that also manifests the anaphoric (retrospective) use, while the proximal term acquires the cataphoric (prospective), typically quotative, use. In Japanese and Finnish, languages with three term distinctions in their systems of demonstratives, it is also the speaker-proximal anaphor (Japanese *koo* ‘like this, in this way’; Finnish *näin* ‘like this’) which is used as a quotative marker.

Based on observations made by Tomasello (1995) and Diessel (2006), attempts have been made to unify deixis and anaphora as a single system of targeting, with the target located near or far in either the speech-external (deictic) or the speech-internal (anaphoric) environment (cf. Talmy 2017). I will not follow this move, since it neglects the parallelism in the contrast between anaphora, the retrospective identification of referents, and cataphora, the prospective identification of referents, and the difference of both from the exophoric (deictic) one.

Demonstratives are an important source for the development of a wide variety of grammatical markers and Brugmann (1904) even assumed that most grammatical categories derive from such expressions. Taking the demonstratives of our focal area as an example, the following overview (Figure 1) can be given for typical exophoric and endophoric uses of MQD-demonstratives and their further developments, on the basis of (a) historical evidence and (b) semantic reconstruction (König 2015; 2017; König & Nishina 2015; König & Vezzosi forthcoming). Detailed empirical evidence for the relevant changes in English and partly also for those in German is provided in König (2017) and in König & Vezzosi (forthcoming). Analogous developments in other languages are partly reconstructed on the basis of these analyses.

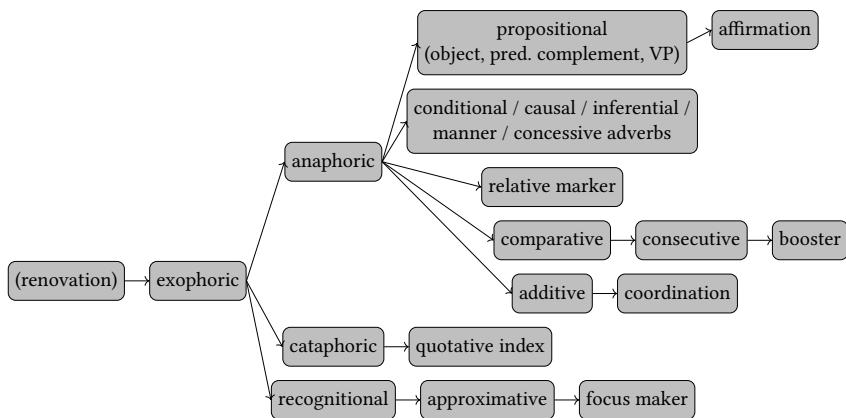


Figure 1: Exophoric and endophoric uses of MQD-demonstratives
(Relative marker, e.g. English *such ... as*; comparative, e.g. English *as ... as*; booster, e.g. English *so good, such kindness!*, Spanish *tan*; additive, e.g. English *also*; coordination, e.g. English *as well as*, French *ainsi que*; quotative index, e.g. English *so to say*, French *ainsi*, German *so*; approximative, e.g. English *ten students or so*, Latin *quasi*)

MQD-demonstratives are an atypical subclass of demonstratives in so far as their exophoric use combined with an appropriate pointing gesture (German *Ich möchte ein solches Fahrrad*. ‘I would like to have a bike like this.’) does not identify the referent directly, but only a type of entity instantiated by the referent (cf. Umbach & Gust 2014; König & Umbach 2018). Figure 1 lists some of the clearly identifiable developments of the Old English demonstrative *swa* (*swa* > *so*; *swalic* > *swelc* > *such*) and the additional developments of *so* which have parallels in many other languages (cf. König 2012; 2015). The examples in (3) illustrate some of these targets of grammaticalisation.

- (3)
 - a. *Will John turn up? – I think/suppose so.*² (propositional object)
 - b. *Such people as John knew did not invite him.* (relative marker)
 - c. *John is not so tall as Bill.* (comparative; degree marker)
 - d. *Bill is also attending the meeting.* (additive focus marker)
 - e. *John’s work has been published in books and journals, as well as in conference proceedings.* (coordinating conjunction for nominals); cf. German *ebenso wie*

²In Japanese it is the hearer-proximal (medial) manner demonstrative *soo* ('like that') which is used as a propositional anaphor.

- f. *Of the 7000 languages or so spoken across the globe the majority is clearly endangered.* (approximative marker)
- g. Spanish/Italian *si*; Polish *tak*; English *yeah swa > yes; quite so* (marker of affirmation)
- h. German *so* (quotative marker): *Die Situation ist schwierig, so die Kanzlerin.* “The situation is difficult,” the Chancellor said.’
- i. *and so on and so forth; and such* (general extender)

In contrast to their counterparts in other Germanic languages, MQD-demonstratives in English (*thus, such, so*) have more or less lost their exophoric use. Complex expressions, separately encoding the deictic and the content dimension (*like this, in this way, therefore*), have taken over that function. Another remarkable feature of the processes of grammaticalisation described in Figure 1 is the fact that these expressions do not manifest the typical co-evolution of meaning and form. Despite a great number of semantic changes there are no concomitant morphological or phonetic changes, which could be due to the stress demonstratives normally carry and the fact that there is very little morphological or phonological substance to begin with. What the relevant examples illustrate, however, is another aspect of grammaticalisation, namely de-categorisation. Grammatical markers and function words derived from MQD-demonstratives manifest an extremely high syntactic versatility and cannot easily be assigned to a specific, limited number of lexical categories.

3 Beyond exophoric and endophoric uses

3.1 Introduction

A detailed look at individual languages and cross-linguistic similarities reveals a variety of other, e.g. discourse-structuring, uses of demonstratives beyond their well-described exophoric and endophoric ones. The major challenge for their analysis is to provide a convincing systematisation and functional analysis for such uses. In the following, four such types of use will be distinguished on the basis of formal and semantic criteria, as discussed in the introduction. Let me add a few remarks as far as the semantic criteria are concerned. The basic semantic structure of demonstratives is a very simple one and comprises two semantic dimensions: (i) a deictic one, identifying a referent in terms of its distance from the centre of orientation (*origo* in the sense of Bühler 1934), that is, as proximal, medial or distal (in languages with a three-term distinction) or proximal

vs. distal (in languages with two-term distinctions) and (ii) an ontological (content) dimension, classifying a referent in terms of such basic semantic notions as entity, object, human being, place, direction, time, manner, quality and degree, among others. Table 1 illustrates this dual semantic structure of demonstratives, frequently also mirrored in a bi-partite morphological structure, with the help of Armenian, a language with a rich and morphologically transparent system of deictic distinctions in contrast to the impoverished systems typically found in West-European languages.

Table 1: System of demonstratives in Armenian

	Definiteness	Entity	Place	Direction
Proximal	ays	sa	aystegh	aystegh
Medial	ayd	da	aydtegh	aydtegh
Distal	ayn	na	ayntegh	ayntegh
	Manner	Quality	Degree	Quantity
Proximal	ayspes	ayspisi	ayschap	aysqan
Medial	aydipes	aydpisi	aydchap	aydqan
Distal	aynpes	aynpisi	ayncchap	aynqan

In some of the uses distinguished in the following sections, either the exophoric (deictic) aspect or the endophoric uses derived from it have either completely disappeared or are no longer clearly visible in English, German, French and Italian. On the other hand, demonstratives may lose the ontological aspect of their meaning or be enriched in various ways by the context and give rise to a variety of adverbial connectives. The resultant semantic distinctions can then be related to several new functions on the basis of their syntactic environments and their meaning. In the following detailed discussion of these uses, I will, largely though not exclusively, use examples and data from the focal area of MQD-demonstratives.

3.2 Coordination of contrasting terms: Loss of deictic component

In coordination of different members of a demonstrative paradigm, the relevant expressions lose their deictic components, while keeping their content component. The relevant sentences simply imply that a situation applies to a whole spectrum of different, non-specific reference points, thus expressing both quantification and vagueness (4)–(5).

- (4) a. English: *here and there, now and then, every now and again, this and that, hither and thither, so so; such and such; neither here nor there* ‘not important, irrelevant’
 - b. German: *so oder so, sowieso* ‘anyway’, *es gibt solche und solche* ‘they come in all colors/kinds’; *hin und her* ‘back and forth, to and fro’, *dies und das* ‘this and that’; *mal so, mal so* ‘this way on one occasion, that way on another’; *dann und wann* ‘now and then’
 - c. French: *ici et là* ‘here and there’; *çà et là* ‘here and there’; *ça se fait comme-çi ou comme-ça* ‘you can do it like this or like that’
 - d. Spanish: *si o asa* ‘like this or like that’; *aquí y allí, aquí y allá* ‘here and there’
 - e. Italian: *qua e là* ‘here and there’; *parlare di questo e quello* ‘to speak about this and that’; *così o cosà* ‘in this or that way, either way, anyway’
- (5) a. *You still find antisemitism here and there.*
 - b. German: (*Im Allgemeinen sind diese Leute tolerant.*) *Aber, es gibt solche und solche.* ‘(In general people are tolerant.) But there are good people and bad people.’
 - c. French (Georges Moustaki, *Ma solitude*): *Elle m'a suivi ça et là, aux quatre coins du monde.* ‘She followed me here and there, all over the world.’
 - d. [A.] *What did you do during your vacation? – [B.] Oh, this and that.*

It is, of course, possible to use such conjunctive combinations of semantically-related demonstratives exophorically, in combination with gestures indicating different locations or instances. This is not their typical use, however, which is meant to withhold precise information.³ Apart from their loss of the deictic component, such coordinate combinations of demonstratives may have some interesting formal properties.⁴ If there is no paradigmatic contrast in the relevant system of demonstratives, one form may be used twice (cf. German *so oder so*)

³It is worth mentioning at this point that coordination of demonstratives from different ontological domains (*here and now; there and then*; German *hin und weg* ‘enthusiastic’) does not have the effect of erasing the deictic meaning.

⁴The loss of the deictic component in these expressions also shows up in the fact that relevant counterparts in other languages do not derive from demonstratives (French *de temps en temps*, German *ab und zu, hin und wieder* ‘now and then’)

or – as examples from Spanish and Italian show (Spanish *así o así*; Italian *così o così*) – a second term may be specifically created for these constructions. In some cases, special unexpected forms are used which manifest a certain phonological and morphological similarity (e.g. French *çà et là*; German *dann und wann*), thus creating a certain formal parallelism. Moreover, subtle semantic distinctions may be found in related pairs. The two French expressions *ici et là* and *çà et là*, which roughly correspond to English *here and there*, in the sense of either location or direction, manifest a subtle difference in meaning, as illustrated by the representative examples in (6) and (7).

(6) French

- a. *J'ai habité ici et là au cours de ma vie.*
‘I have lived here and there in the course of my life.’
- b. *J'aime me promener ici et là le dimanche.*
‘I love walking here and there on Sundays.’

(7) French

- a. *Mon fils a déposé ses affaires ça et là dans la maison.*
‘My son put his things all over our house’
- b. *J'ai trouvé des réponses ça et là dans ma mémoire.*
‘I have found answers here and there in my memory’

The difference between these minimal pairs seems to be one of identifiability, which is still possible with the first expression, but not with the second, where the motion is less orderly and more chaotic.⁵ A similar distinction can be found between *to and fro* vs. *hither and thither* in English, where the old-fashioned expressions suggests more chaotic motion.

In many European languages, these binomials typically have the proximal expression in first position, but this may differ from language to language. In Japanese, a language with pervasive three-term deictic distinctions, the medial or distal member of the paradigm invariably precedes the proximal one, as illustrated by examples in (8). Note that the idiomatic English translations reverse the order of the relevant expressions in Japanese, e.g. *soko-koko* ‘lit. there-here’.

(8) Japanese : *soko-koko* ‘here and there’; *atti-kotti* ‘this way and that way, to and fro’; *are-kore* ‘this and that’; *sonna-konna* ‘like this and like that’; *soo-koo* ‘either way, anyway’

⁵This observation is based on consultation with native speakers of French (Bernard Faurie and Claire Moyse).

3.3 Idiomatic use as discourse-structuring device: Loss of deictic and ontological components

It is a well-known fact that demonstratives may lose their deictic meaning, and even the other, ontological, aspect of their original meaning as well, by developing into grammatical markers. The existential use of English *there* is a case in point. Originally used on the basis of the general principle that existence implies “being in some location” (cf. Lyons 1967), the local demonstrative lost both its deictic (distal) meaning and its locative meaning. Both uses, the basic locative one and the derived existential one, are visible in (9).

- (9) *There is a dog over there.*

On the other hand, a demonstrative may lose only its deictic component and develop a purely lexical meaning on the basis of its ontological component. The German counterpart of the local demonstrative *there* in English, namely *da*, is a clear example of such a development, denoting as it does the destination of a journey, walk or motion in general, as well as an important subset of relevant destinations, namely one’s home (10).

- (10) German
- a. Exophoric use
Das Haus, das wir suchen, ist da drüber.
‘The house that we are looking for is over there.’
 - b. Non-deictic use⁶
In einer Stunde sind wir da.
‘We will have reached our destination in an hour.’
 - c. Non-deictic, non-anaphoric use
Karl ist nicht da.
‘Karl is not home.’

The use to be discussed in what follows is a more general phenomenon and shows up in a variety of frozen idioms. In the following examples the demonstratives have all lost their deictic components, and partly also their content components. They are found in short idiomatic utterances which are typically used at transition points of verbal interactions. Given that the communicative sense of such expressions cannot easily be translated into another language, I will mainly

⁶Note that English *there* can also have this function: *We are almost there.*

discuss examples from two genealogically close languages, namely English and German.

Before these examples are analysed in detail, a quick word on the collection of data for this analysis is required. While the examples used so far are clear instances of grammatical constructions in English, neither requiring consultation with native speakers nor any documentation by relevant corpora, the following examples are retrievable from dictionaries or usage manuals. An analysis of their use and function would ideally have to be based on relevant examples with a rich contextual embedding, but the use of such data would require far too much space⁷ given that in corpus-based studies of such expressions only a single expression is typically discussed in detail at a time (cf. Golato & Betz 2008; Barske & Golato 2010). For reasons of space and my goal of assessing how much of their ontological, deictic, anaphoric and cataphoric meaning these uses have lost and kept, I will therefore partly rely on such detailed studies whenever they are available, but will have to use information provided by dictionaries and introspection in many other cases.

Instead of presenting relevant data in an unordered list, I will divide them into two groups depending on whether they are primarily used in introductory, initiating conversational moves (11) or in responsive, conclusive ones (12).

- (11) a. *Now then!*
(introductory move, initiating, attention getter and topic change)

b. *Here goes!* (introduction to doing something brave, risky, foolhardy)

c. *Here we go.* (introductory, exhortative)

d. *Hi there!* (introductory, greeting)

e. *There you are. There you go.* (introductory, offer, providing service)

(12) a. *So there!*
(responsive, closing an argument by maintaining a decision or view)

b. *And that's that.*
(responsive, closing an argument, terminating further discussion)

c. *That's it!* (responsive, closing an argument through agreement)

d. *Here we go again.*
(responsive, comment on the recurrence of an unpleasant situation)

⁷To illustrate the enormity of such a task, let me just mention that for the use of the English discourse marker *well* alone there are at least ten detailed studies.

- e. *Could you pass me the sugar? – Here you go.*
(responsive, compliance with a request)
- f. *There you go.* (responsive, accepting an unsatisfactory situation)
- g. *There we were/there he was.*
(closing, summing up, slowing down a story)
- h. *Now you're talking.* (responsive, confirming the relevance of the preceding turn for an argument)
- i. *Don't lose them now. They are my favorite gloves.*
(responsive, emphatic injunction)
- j. *There is a good boy!* (responsive, approval or encouragement)
- k. *There, there!* (responsive, attempt to comfort someone)
- l. *Now really!* (reprimanding)
- m. *Come now!* ‘Don't exaggerate!’

Whether the development of such uses of demonstratives should be regarded as an instance of grammaticalisation or instead as an instance of “pragmaticisation” is a matter of some controversial debate (cf. Diewald 2011; Wiese 2011; Degand & Evers-Vermeul 2015). It is true that these examples do not manifest the co-evolution between meaning and form (formal attrition + semantic bleaching, etc.) that we find in prototypical instances of grammaticalisation, but then – as noted above – there are very few formal reductions and changes observable in the historical development of nearly all demonstratives. What the uses illustrated in (11) and (12) certainly manifest are extensions to new contexts, persistence, de-categorisation and semantic bleaching. In view of these changes I will regard the use of demonstratives in (11) and (12) also as a result of grammaticalisation, especially since a concept of grammar that excludes dialogue-structuring devices would be too restrictive (cf. Degand & Evers-Vermeul 2015: 74).

In contrast to the heterogeneous set of expressions frequently subsumed under the term “discourse markers”, the expressions under discussion manifest a general property that justifies their inclusion under a common term. As already pointed out, these expressions typically occur at transition points of a verbal interactions, i.e. at points typically involving a change of speaker, topic, or some subpart of an interaction. This characterisation classifies these utterance types in terms of a more precise discourse function, that of being responsive and possibly terminating a phase of conversational interaction or of being a first move, an initiative, in a new subpart of a verbal exchange. It is this occurrence at transition

points that justifies the label “discourse-structuring devices”.⁸ Marking the end of an exchange can also be a signal for the beginning of a new part and some of the expressions listed above can mark both the end of a part of an interaction and the beginning of a new one.⁹

The retrospective versus prospective orientation of these expressions is, of course, reminiscent of the contrast in the standard endophoric uses between anaphoric and cataphoric uses. Indeed, these orientations can be regarded as the last trace of a more basic endophoric use of demonstratives preserved in these expressions. The typical features of anaphora and cataphora, however, are no longer there. There is no referent identified via a preceding or following expression, but simply a look back or forward to some part of an interaction.

Analogous expressions with MQD-demonstratives are also found in German. One can never be sure that one will catch all the fish in a certain terminological net, but it seems that the inventory of such expressions is smaller in German than in English and is primarily based on further developments of the manner demonstrative *so* (13).

(13) German

- a. *Ach so.* ‘O.k., I understand.’ (responsive, removing epistemic asymmetry)
- b. *So, so!* (responsive and initiating, accepting a previous statement and its implications)
- c. *So! So, das wär’s für heute!* ‘Right/well/so, that’s it for today!’ (responsive, boundary signal after an “accomplishment”)
- d. *Also!* ‘Now then!’ (initiating)
- e. *Also dann!* (responsive, conclusive)
- f. *Na so was!* ‘Well I never!’ (responsive, comment expressing surprise)
- g. *Na also!* (responsive, accepting a repeated attempt as a solution)
- h. *Nun denn.* ‘Well, all right.’ (responsive, signalling reluctant agreement)

⁸Some of the examples in (11) and (12) have also been subsumed and discussed under the term *interjections*. While the use of this term may have some justification for a characterisation of the distribution of the expressions under discussion, it does not say very much about their function.

⁹Similar extended uses of the basic manner demonstratives *nii* and *soo* in Estonian are described in Keevallik (2005; 2010), where these two “pro-adverbs of manner” and their counterparts in several other languages are analysed as marking transitions between conversational activities.

- i. *Nun, wir wissen das nicht genau.* ‘Well, we don’t know for sure’¹⁰ (responsive, not providing a straight answer)
- j. *Warum willst du dahin gehen? – Nur so.* ‘Why do you want to go there? – I simply do.’ (responsive, meant to indicate the lack of a reason for doing something)
- k. *Plötzlich schrie er mich an. Einfach so.* ‘Suddenly he was shouting at me. Just like that.’ (initiating, indicating that somebody had no reason for doing something unusual or outrageous)

In French, there seem to be fewer discourse-structuring devices derived from demonstratives, the distal demonstrative of time *alors* (< (à +) Latin *illa hora* ‘at that hour’) being the most frequent and versatile expression in this use (*et alors* ‘so what?’; *alors* ‘and?’; *ça alors* ‘my goodness/well, really!'; *merde alors* ‘holy shit’, ‘I’ll be damned’; *mince alors* ‘wow/blimey/for heaven’s sake’). As shown by the English glosses in the examples above, *alors* is primarily used as a responsive discourse-structuring device, but can also indicate a change of topic (cf. Degand & Fagard 2011). In contrast to the distal demonstrative *alors*, which derives from a demonstrative in Late Latin (à *illa hora*), its proximal counterpart *maintenant* has no deictic origin and is one of the rare cases which derive from a construction composed of major lexical elements, analogous to *stante pede* in Latin.

To summarise, idiomatic expressions based on demonstratives may have a use as speech acts (*here is to you!*; *now, now!*; *hi there!*), but are more typically used as discourse-structuring devices and relate to overarching purposes and intentions of speakers and goals of an argument or discourse. There are, of course, striking differences in the form and meaning of such expressions within a language and across languages, but a common feature can be seen in the discourse-structuring function of either closing or opening a sub-sequence of an argument or interaction. These functions are clearly related to the anaphoric and cataphoric uses of the relevant demonstratives, but have to be kept distinct from these more basic uses of the same expressions. The examples in (11) and (12) suggest that local and temporal demonstratives are most frequently used for such idiomatic extensions in English, whereas it is the manner demonstrative *so* which typically gives rise to such discourse-structuring devices in German.

¹⁰In Jucker (1993: 450), the function of *well* is described as indicating a shift in the relevant context.

3.4 Reference via different points of orientation: Loss of deictic component

The characteristic feature of the third type of special (non-canonical or extended) uses of demonstratives is a transfer of the point of orientation (*origo*) from the coordinates of an utterance or a point in the development of a text to some point provided by norms, expectations, evaluations, and so on, of speakers. The relevant use thus involves an additional subjective implication and can therefore also be considered as an instance of Traugott's general tendency of semantic change from propositional to expressive meaning or, more specifically, from meanings based in the external described situation to meanings based in the internal (evaluative / perceptual / cognitive) described situation (Traugott 1982; 1990). The English (partly defective) demonstrative of degree *so*, which can still be used exophorically, provides particularly good illustration of this shift in the *origo*, which could also be described as a loss of the standard anchoring points of a deictic dimension. Consider the examples in (14).

- (14) a. *The fish I caught yesterday was so* (+ gesture) *long.*
(exophoric, gestural use)
- b. *Why don't you do it like so* (+ gesture)?
(exophoric, gestural use)
- c. *My fish is not so long/as long as the one you caught.*
(standard of comparison)
- d. *Her hair always has to be just so.*
(point of orientation provided by norm)
- e. *You can only eat so much.*
(point of orientation provided by capacity)
- f. *Mrs. Jones is always so/soooo helpful.*
(point of orientation provided by expectation)
- g. *I am so taking a long nap today.*
(point of orientation provided by contradictory view)

In (14a) and (14b), the point of orientation is provided by the speech situation and a pointing or mimicking gesture. This exophoric use of *so*, denoting manner or degree has become marginal in present-day English, *like this* and *that* being typically used instead, cf. *Why don't you do it like this?* and *The fish I caught yesterday was so/that long.* The other uses of *so* are not exophoric ones. The point of orientation in (14c) is provided by the standard of comparison in a comparative construction. In (14d) it is a normative one ('as it should be') and in (14e) a

point is indicated that denotes general limits of a specific action, that of eating in this case. The example in (14f) is generally analysed as an exclamation and identifies a point surpassing a certain standard as point of orientation. Together with *very*, *so* is among the five most frequently used intensifiers in English and characterises a point on a scale evaluated by the speaker as high or beyond expectation (cf. Bolinger 1972). This use as an intensifier or booster can also be found in the use of degree demonstratives in other Germanic languages, cf. German *Sie ist so intelligent!* and Dutch *Ze is zo intelligent!* ‘She is so intelligent!’. Finally, (14g) is an instance of a recent development in American English, the relevant use of *so* being generally referred to as “pre-verbal use” (Boulonnais 2005) or as “generation-X” *so* (Zwicky 2007). In that use, *so* invariably bears a focal stress and emphasises the truth or factuality of a statement in the context of a contradictory view. The underlying points of orientation of the first five uses of the degree adverb *so* can roughly be illustrated by Figure 2, which is simply meant to indicate that different points of orientation are involved:

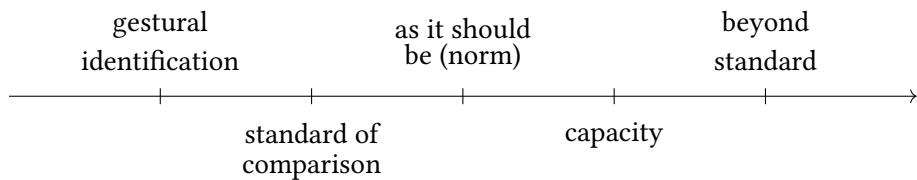


Figure 2: Degree-denoting uses of *so* in English

How can these extended uses of the demonstrative degree marker *swa* (> *so*) be analysed and explained? Note first of all that the last four uses are specific phenomena of English; only the exophoric uses (14a)–(14b) and the comparative endophoric use (14c) have clear parallels in other Germanic languages. Since an analysis of the last four uses as endophoric, i.e. as relating to a portion of preceding or following context, is excluded, a possible analysis might be to analyse them as directly deriving from the exophoric use and a shift of the *origo* from the speech situation to some evaluative dimension. Extended evaluative uses of demonstratives are a widespread phenomenon, also showing up in the use of *this* ('positive') and *that* ('negative') in English (*I like this.* vs. *Oh, that.*). If such salient evaluations are frequently used in specific contexts (e.g. *just*, certain types of verbs, contexts of contradicting) they can become established as being part of certain constructions.

3.5 Connectives expressing rhetorical relations: Loss of deictic component and enrichment of meaning

Our fourth type of non-canonical uses of demonstratives is characterised by a loss of the deictic component, combined with an enrichment, rather than a loss, of ontological meaning. A typological study of such developments is presented in Diessel & Breunisse (this volume). Again using examples from the focal area of MQD-demonstratives, the adverbial use of English *so* provides rich resources for this discussion. The examples in (15) illustrate the ability of English *so* (< Old English *swa*) to express a variety of rhetorical relations.

- (15) a. *I did not like it. So I wrote to him.* (causal)
b. *The whole thing was tied up in knots, so that we were not able to undo it.* (resultative)
c. *He went into lower gear, so (that) his car would slow down.* (purposive)
d. *He is very sick. Even so he goes to work.* (concessive)
e. *So you are a linguist, eh? - So what?* (inferential)

How can these uses be best characterised? Since *so* does not relate to preceding utterances, but to states of affairs or propositions, the above examples are not instances of textual deixis. Their retrospective perspective certainly gives them an anaphoric character, but their semantic enrichment through the accompanying co-text clearly differentiates them from the anaphoric use as VP-anaphora and propositional anaphora found in examples like (16).

- (16) a. *Bill writes his essays in the library and Mary does so at home.*
b. A. *Italy is in a very difficult situation at the moment.* –
B. *I think/suppose so.*
c. *It might rain tomorrow. If so we will postpone our trip.*

Another question requiring a clear answer concerns the decision between polysemy and vagueness for the uses in (15). I will opt for vagueness for the following reasons: all uses of *so* in (15) are based on the manner reading of an anaphorically-used demonstrative. The more specific interpretation in the context of these anaphors can be derived from this manner component and the contextual information provided in the rest of the sentence. The crucial context in (15d) is the focus marker *even*, which in combination with *so*, in contrast to conditional *if* (*even if p, q*), results in a factual, concessive interpretation. The crucial contextual information in (15a)–(15c) is the preceding factual assertion, which

enriches the manner interpretation to a causal one in (15a) and in combination with the complementiser *that* to an interpretation as consequence, either factual (resultative) or as an intention (purposive) in combination with hypothetical modality (15c). Finally, it is a situational context of partial ignorance and the interrogative character of the utterance that results in the overall meaning ‘inference’ in (15e).

Similar uses to (15) can be found for *so* in German, for *così* in Italian and for *ainsi* in French. One of the rhetorical relations in the preceding list of examples that is notably absent in English is conditionality. The conditional use of the connective *so* is somewhat archaic in Modern English and expresses a necessary condition, in contrast to the unrestricted use of German *so* and French *si/ainsi* in both the protasis and the apodosis of all types of conditionals (17).

- (17) a. German: *So(fern) er rechtzeitig kommt, können wir in die Oper gehen.*
‘We can go to the opera, provided he comes in time.’
- b. German: *Kommt er rechtzeitig, so können wir in die Oper gehen.* ‘If he comes in time we can go to the opera.’
- c. German: *So Gott will, können wir morgen nach Italien fahren.* ‘God willing, we can go to Italy tomorrow.’ (conditional)
- d. English (OED): *It is no matter how dirty a bag it is conveyed to him in, he will accept it, so the money is good.* (necessary condition)
- e. French (cf. Italian *se*, Spanish *si*): *Si Pierre venait aussi, (ainsi) nous pourrions jouer au tennis.* ‘If Pierre came too, we could play tennis together.’

It is a characteristic feature of German *so* that this demonstrative combines with several adverbs, adjective and prepositions to derive subordinating conjunctions (18a) and adverbs (18b). The relevant English counterparts of the conjunctions in German often take the form of comparative constructions.

- (18) German
 - a. *sofern* ‘providing’, *sobald* ‘as soon as’, *soweit*, *so viel* ‘as far as’, *so lange* ‘as long as’
 - b. *sowohl* ‘both/as well’, *sofort* ‘immediately’, *sogleich* ‘at once’, *somit* ‘consequently, therefore’, *sowie* ‘as well as’, *sodann* ‘thereupon, then’

If we go beyond our focal area, English *hence*, *thence*, *therefore*, *thereby*, *hereby*, *but then*, *thus* also have to be included in the use type under discussion. The

token-reflexive use of *hereby* in combination with a performative prefix (e.g. *I hereby declare ...*) identifies the following proposition as an explicit illocutionary act. Its distal counterpart *thereby* lost its exophoric use and has a resultative meaning in a basically anaphoric use (19).

- (19) a. *I hereby promise you never to touch hard drinks again.*
b. *John knocked over the red wine, thereby ruining the table cloth.*

Up to the time of Early Modern English, the system of demonstratives in that language also included directional demonstratives, a proximal pair (*hither, hence*) formally related to the proximal locative expression *here*, and a distal pair (*thither, thence*) analogously related to the distal locative expression *there*. The first member of both pairs was used to signal movement towards a goal, the location of the speaker in the case of *hither* and a contextually distant location the case of *thither*. The second member signalled motion away from some source, from the *origo* in the case of *hence* and from some contextually given place in the case of *thence*. Just like most MQD-demonstratives in English, these directional uses became archaic and thus highly formal in their exophoric use (*hither*) or lost that use completely (*hence*). In Modern English, *hence* is primarily used in an argumentative or rhetorical sense and introduces an inference based on a preceding premise (20).

- (20) *John got a pay rise, hence the new car.*

4 Conclusion

Based on general typological surveys and on previous studies of MQD-demonstratives, this chapter has shown that such demonstratives have acquired a variety of additional uses as a result of losing their deictic meaning, and partly also their ontological meaning, and have developed into specific constructions, grammatical markers, or discourse-structuring devices. On the basis of data from a small subset of demonstratives, the analysis of the second type of extension beyond anaphoric and cataphoric uses was the main focus of the chapter. Four types of extended, non-canonical uses were discussed: (a) coordination of demonstratives, typically of expressions denoting the same ontological aspect but distinct deictic values, (b) short, frozen idiomatic constructions without lexical content, whose basic endophoric use is still visible in their introductory or responsive functions, (c) extended exophoric uses of the former demonstrative *so*

(< Old English *swa*) with a shift of the centre of orientation from the speech situation to a point determined by some evaluation, and (d) basically anaphoric uses of demonstratives, metonymically enriched with meaning relating to rhetorical structure. Instead of assigning the relevant changes to a process of pragmaticisation, essentially defined as a loss of truth-conditional content, I followed the view expressed *inter alia* in Degand & Evers-Vermeul (2015) that these changes manifest crucial properties identified for grammaticalisation and can therefore be regarded as instances of such well-known changes.

It was also shown that, in English, demonstratives from several ontological dimensions have developed such extended uses as discussed above, while in German MQD-demonstratives play a prominent role in such developments. Both English and German are languages with impoverished systems of demonstratives, both in the differentiation of deictic distinctions and in the relevant ontological distinctions, but both have developed a variety of extended uses beyond the exophoric and endophoric ones. Developments of this kind, starting out from demonstratives and resulting in discourse-structuring expressions or discourse markers, are a wide-spread phenomenon. In Auer & Maschler (2016) it is shown that such changes starting out from the common source of proximal temporal demonstratives (*nu/nå*) are a characteristic areal feature of European languages.

Acknowledgements

I am indebted to two anonymous reviewers, to Volker Gast and to Åshild Næss, as well as to the participants of two conferences on demonstratives (Oslo, June 14–15, 2018; Rijeka, June 28–29, 2019) for pointing out various inadequacies and suggesting possible improvements. All remaining errors are my own responsibility.

References

- Arbib, Michael A. 2012. *How the brain got language: The mirror system hypothesis*. Oxford: Oxford University Press.
- Auer, Peter & Yael Maschler (eds.). 2016. *NU/NÅ: A family of discourse markers across the languages of Europe and beyond*. Berlin: De Gruyter.
- Barske, Tobias & Andrea Golato. 2010. German *so*: Managing sequence and action. *Text & Talk* 30(3). 245–266.
- Bolinger, Dwight. 1972. *Degree words*. The Hague: Mouton.

- Boulonnais, Dominique. 2005. So et les phénomènes d'expansion contextuelle. In Geneviève Girard-Gillet (ed.), *Aux marges du texte: Texte et co-texte* (CIEREC Travaux 128), 49–69. Saint-Étienne: Publications de l'Université de Saint-Étienne.
- Brugmann, Karl. 1904. *Die Demonstrativpronomina der indogermanischen Sprachen*. Leipzig: Teubner.
- Böhler, Karl. 1934. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Fischer.
- Degand, Liesbeth & Jacqueline Evers-Vermeul. 2015. Grammaticalization or pragmaticalization of discourse markers? More than a terminological issue. *Journal of Historical Pragmatics* 16(1). 59–85.
- Degand, Liesbeth & Benjamin Fagard. 2011. Alors between discourse and grammar. The role of syntactic position. *Functions of Language* 18(1). 29–56.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Diewald, Gabriele. 2011. Pragmaticalization (defined) as grammaticalization of discourse functions. *Linguistics* 49(2). 365–390.
- Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1). 61–112.
- Gerner, Matthias. 2009. Deictic features of demonstratives: A typological survey with special reference to the Miao group. *The Canadian Journal of Linguistics / La Revue Canadienne de Linguistique* 54(1). 43–90.
- Golato, Andrea & Emma Betz. 2008. German *ach* and *ach so* in repair uptake: Resources to sustain or remove epistemic asymmetry. *Zeitschrift für Sprachwissenschaft* 27. 7–37.
- Jucker, Andreas H. 1993. The discourse marker *well*: A relevance theoretic account. *Journal of Pragmatics* 19. 435–452.
- Keevallik, Leelo. 2005. The deictic *nii* ‘so, in this way’ in interaction. In Daniele Monticelli, Renate Pajusalu & Anu Treikelder (eds.), *From utterance to uttering and vice versa: Multidisciplinary views on deixis*, 109–126. Tartu: Tartu Ülikooli Kirjastus.
- Keevallik, Leelo. 2010. Pro-adverbs of manner as markers of activity transition. *Studies in Language* 34(2). 350–381.
- Kibrik, Andrej A. 2011. *Reference in discourse*. Oxford: Oxford University Press.

- König, Ekkehard. 2012. Le rôle des déictiques de manière dans le cadre d'une typologie de la déixis. *Bulletin de la Société de Linguistique de Paris* 107(1). 11–42.
- König, Ekkehard. 2015. Manner deixis as source of grammatical markers in Indo-European languages. In Carlotta Viti (ed.), *Perspectives on historical syntax*, 33–60. Amsterdam: Benjamins.
- König, Ekkehard. 2017. The deictic identification of similarity. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 143–164. Amsterdam: Benjamins.
- König, Ekkehard & Yoko Nishina. 2015. Deixis der Art und Weise, der Qualität und des Grades im Deutschen und Japanischen: Eine kontrastiv vergleichende Analyse. In Yoko Nishina (ed.), *Sprachwissenschaft des Japanischen* (Linguistische Berichte Sonderheft, 20.), 7–31. Hamburg: Buske.
- König, Ekkehard & Carla Umbach. 2018. Demonstratives of manner, of quality and of degree: A neglected subclass. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives: Syntax, semantics and pragmatics*, 285–328. Berlin: De Gruyter Mouton.
- König, Ekkehard & Letizia Vezzosi. Forthcoming. From OE *swa* to ModE *so*: Similarity in manner, quality and degree as a basis of grammatical markers. In Bettelou Los (ed.), *English historical linguistics 2018*. Amsterdam: Benjamins.
- Lyons, John. 1967. A note on possessive, existential, and locative sentences. *Foundations of Language* 3(4). 390–396.
- Talmy, Leonard. 2017. *The targeting system of language*. Cambridge, MA: MIT Press.
- Tomasello, Michael. 1995. Joint attention as social cognition. In Chris Moore & Philip J. Dunham (eds.), *Joint attention: Its origin and role in development*, 103–130. Hillsdale, NJ: Erlbaum.
- Traugott, Elizabeth Closs. 1982. From propositional to textual to expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics*, 245–271. Amsterdam: Benjamins.
- Traugott, Elizabeth Closs. 1990. From less to more situated in language: The unidirectionality of semantic change. In Sylvia M. Adamson, Vivien A. Law, Nigel Vincent & Susan Wright (eds.), *Papers from the Fifth International Conference on English Historical Linguistics*, 497–518. Amsterdam: Benjamins.
- Umbach, Carla & Helmar Gust. 2014. Similarity demonstratives. *Lingua* 149. 74–93.
- Wiese, Heike. 2011. So as a focus marker in German. *Linguistics* 49(5). 991–1039.

Ekkehard König

Zwicky, Arnold M. 2007. *GenX* so. <https://arnoldzwicky.org/2011/11/14/genx-so>.

Chapter 3

The use of manner demonstratives in discourse: A contrastive study of Wan (Mande) and Kambaata (Cushitic)

Tatiana Nikitina

LLACAN (CNRS, INALCO)

Yvonne Treis

LLACAN (CNRS, INALCO)

This chapter compares manner demonstratives in two unrelated African languages, Kambaata (Cushitic, Ethiopia) and Wan (Mande, Côte d'Ivoire). Both languages have specialised manner demonstratives, yet differ strikingly in their typological profile and in the way the manner demonstratives behave syntactically. Through systematic comparison of data from both languages, similarities, which are likely due to common semantic mechanisms of meaning extension, and differences, which are likely due to structural differences between the languages, are identified. It is argued that, despite the shared core meanings, manner demonstratives belong to different syntactic classes in Kambaata and in Wan. The difference in syntactic category helps account for the striking dissimilarities in the range of attested extended uses.

1 Introduction

Manner demonstratives are deictic expressions that identify a way of carrying out an event or the extent to which a property holds, as observed in the speech situation (exophoric use) or expressed in the preceding discourse (endophoric anaphoric use). Manner demonstratives are generally assumed to be adverbs (cf. Diessel 1999: 74). We take as our starting point a semantic definition of “manner demonstrative” since, as we will see, the syntactic category to which they belong



in different languages may vary. Despite a recent increase of interest in manner demonstratives on the part of the typological community (e.g. Guérin 2015, König 2017, König & Umbach 2018), they remain poorly studied even in well-described languages. The ways they function in discourse, in particular, remain critically underexplored, so that the few existing studies focusing on their grammaticalisation (e.g. König 2015) are still dissociated from synchronic corpus-based studies of new and emergent usage.¹

This study is a contrastive description of the use of manner demonstratives in two unrelated African languages, Kambaata (Cushitic, Ethiopia) and Wan (Mande, Côte d'Ivoire). We chose to compare these two languages, since they both have specialised manner demonstratives, yet differ strikingly in their typological profile and in the way the manner demonstratives behave syntactically. Through systematic comparison of data from Kambaata and Wan we hope to identify similarities (which are likely due to common semantic mechanisms of meaning extension) and differences (which are likely due to structural differences between the languages), and to address the challenge of accounting for these similarities and differences within the same model.

We first discuss and classify the common uses attested in both languages (§2); these uses, we argue, correspond to the “core” meanings of manner demonstratives, which we predict to cluster in other languages as well. We then turn to differences in the ways manner demonstratives behave syntactically in the two languages. We argue that despite their close semantic similarity, manner demonstratives belong to different syntactic classes in Kambaata and in Wan (§3). The difference in syntactic category helps us account for the striking dissimilarities in the range of attested extended uses. We use a semantic map approach to model the differences, and end with a brief discussion of the study’s methodological and theoretical implications (§4).

2 Same core uses

2.1 Contextually salient manner and represented speech

Manner demonstratives of Kambaata and Wan are characterised by strikingly similar core uses. They occur most frequently in two types of context, indexing either contextually salient manner or instances of represented speech.

¹Notable exceptions include König & Nishina (2015) on Japanese, Shor (2018) on Hebrew *kaxa*, Karssenberg & Lahousse (2018) on French *ainsi*, and Keevallik (e.g. 2005, 2010) on Estonian.

Contextually salient manner can be inherent in the current speech situation, as in (1a) and (1b), or it may be suggested by an accompanying gesture, as in (2a) and (2b). Both uses are deictic: in order to interpret the manner to which the demonstrative refers, it is important to observe the situation in which they were uttered. In (1a), the manner demonstrative refers to the way the addressee is acting. In (1b), it indexes the looks of the child present at the site where the sentence is uttered.²

(1) Manner inherent in the current speech situation

a. Kambaata [written]

Lankíi harabas-á **hitt-ítá** torr-itókkoont!

again dirt-M.ACC like.this[1]-F.ACC throw-2F.APPR

(Speaker sees addressee dropping dirt in the front yard and tells her off:) ‘Don’t throw the dirt away again **like this!**’

b. Wan

àà bāā n̄é é kēé wà, bāā n̄é é zièziè yā
INTJ LOG+ALN child DEF like.this not LOG+ALN child DEF ugly with
kēé wà
like.this not

‘Oh, my child is not like this, my child is not ugly **like this.**’

In (2a) and (2b), the relevant manner is represented by a gesture, so that, in order to interpret the manner demonstrative, the listener needs access to the accompanying visual information.

(2) Manner suggested by an accompanying gesture

a. Kambaata [elicited]

Hitt-ítá ass-í fann-óomm

like.this[1]-F.ACC do-1SG.PFV.CV open-1SG.PFV

(Speaker demonstrates his opening technique:) ‘I opened it **like this.**’

²All Wan examples are from a corpus of spontaneous oral data (consisting mostly of narratives). The Kambaata examples are from three different types of sources: spontaneous oral data (indicated as: [oral]), local written publications (indicated as: [written]) and from elicitation (indicated as: [elicited]). In the data collection, direct translation elicitation was generally avoided. Instead speakers were asked to come up with near-natural mock dialogues for situations that were laid out by the researcher. Additional elicitations were made on the basis of examples attested in recordings, written text data and mock dialogues.

b. Wan

è é ɔ wō kēé
3SG REFL hand did like.this

'He made a gesture with his hand like this.'

In addition to pointing to a contextually salient manner, the same markers can also be used to refer to a following instance of reported speech. Unlike in (1) and (2), the context in which the manner demonstrative is interpreted in (3) is non-concomitant with the demonstrative's utterance. Unlike the accompanying gesture, representation of speech necessarily follows the use of the demonstrative in time, so this use could be treated, strictly speaking, as discourse-cataphoric. Yet that difference seems to derive directly from the fact that gestural representation relies on the multimodal potential of oral discourse, which is unavailable in the case of representation of speech, for rather technical reasons.

(3) Reference to represented speech

a. Kambaata [written]

Ká-s haar-óo xah-á
A_DEM1.M.ACC-DEF new-M.ACC thing-M.ACC
dagg-oommí-i shool-kí bar-é gassim-á
know-1SG.PFV.REL-NMZ1.M.NOM four-ORD day-M.GEN morning-M.ACC
ées hitt-íta y-itoonté-'e-ta
1SG.ACC like.this[1]-F.ACC say-2SG.PFV-1SG.OBJ.REL-F.COP2
j-éechch-u ...
time-SG-M.PRED

'I came to know that new detail on the morning of the fourth day,
when you said to me like this ... [followed by a direct quote].'

b. Wan

bé mójj zō bé mōjj gé é dè lèjj kēé bé à
then LOG.PL came then LOG.PL said REFL father to like.this then 3SG
dè gé àjj lèjj kēé ...
father said 3PL to like.this

'And they came to their father and said like this, and their father said
like this ... [followed by a representation of their interaction with the
father, by means of a dialogue].'

Both the salient manner and the represented speech use have discourse-anaphoric equivalents, where the demonstrative refers to a description in the preceding discourse, rather than to the concomitant situation or a following speech

representation. In (4), the relevant manner is suggested by a preceding description; the description can be concise or potentially comprise an entire portion of the preceding narrative.

(4) Anaphoric reference to previously described manner

a. Kambaata [oral]

Tah-íchch-u dágó hitt-ítá afuu'll-ít
 fly-SGV-M.NOM suddenly like.this[1]-F.ACC sit-3F.PFV.CV
 zug-gáni-yan waall-ó=da ...
 lie.in.ambush-3F.IPfv.CV-DS come-3M.PFV.REL=COND

'When a fly comes suddenly while it (lit. she = the chameleon) is lying (lit. sitting) in ambush **like this** (= in a way previously described, i.e. without moving, apart from its eyes going up, down, left, right) ...'

b. Wan

bé àjì màjì yì é bō, kéké dí kéké lèjì kéké ...
 then 3PL rice water DEF served if 3PL+3SG offered that to like.this
 'And they served the rice water, and when they offered it to someone **like this** (= in a way previously described, i.e. boiling hot, brought directly to the lips) ...'

The demonstratives can also refer back to a preceding representation of speech, as in (5).

(5) Anaphoric reference to preceding representation of speech

a. Kambaata [written]

... ées áchche hittig-úta y-ee-'é
 1SG.ACC then like.this[2]-F.ACC say-1SG.PFV-1SG.OBJ.REL
 j-áata iyy-áqq-u' góoff
 time-F.ACC carry-MID-M.NOM-1SG.POSS finish.3M.PFV.CV
 fájj-o
 do.completely-3M.PFV

'Then when he said **like this** (= as quoted in the preceding discourse) to me, ... my patience was exhausted.'

b. Wan

bé kólé né é wō é gbjé lè kéké
 then man DIMIN DEF did REFL manner on like.this

'And the old man acted in this way **like this** (= as she has said).'

As seen in the preceding examples, Kambaata has two manner demonstratives, *hittíta* and *hittigúta*, glossed ‘like.this[1]’ (see 1a, 2a, 3a, 4a) and ‘like.this[2]’ (see 5a), respectively. Both are morphologically transparent: they consist of a simple or extended demonstrative base *hitt-* / *hittig-* plus a portmanteau morpheme of feminine gender and an adverbial case, i.e. either the accusative,³ the instrumental or the oblique case (see Treis 2019 for details on the morphology). Diachronically, the extended stem *hittig-* is the result of a merger of a demonstrative and a similative morpheme *-g ‘like’. The two manner demonstratives, *hittíta* and *hittigúta*, are interchangeable in the context of the core uses described above.

Table 1 summarises the uses of the manner demonstratives described in this section, with reference to the relevant examples. All four types of use are widely attested in Kambaata and in Wan, forming the “core” of the manner demonstrative category. The close similarity of these uses in two unrelated languages suggests that the category may be cross-linguistically relevant.

Table 1: Core uses of manner demonstratives in Kambaata and in Wan

	Grounded in the current speech situation	Anaphoric
Contextually salient use	(1), (2)	(4)
Reference to represented speech	(3)	(5)

2.2 The ‘just so’ implicature

Besides the core functions discussed in §2.1, both languages display a surprisingly similar range of uses that we believe is derived by a common mechanism of Gricean implicature (Grice 1975). Suppose that the manner demonstrative is uttered in the absence of any contextually salient indication of manner. If the listener assumes that the speaker is being cooperative, she would be led to believe that by referring to manner when no particular circumstances are suggested by context, the speaker implies that the event involved *no* circumstances that are worth mentioning. This could imply, for example, that an action that is usually performed for a particular reason was in this particular case performed for no

³The accusative case is the case of direct objects but also of adverbial constituents in Kambaata, see e.g. (3a) in which *xah-á* ‘thing’ is the accusative-marked direct object of *dag-* ‘know’ and *gassim-á* ‘morning’ an accusative-marked noun in adverbial function.

reason at all ('just so'), or perhaps an event normally involving a long preparatory stage in this particular case happened without any preparation.

The 'just so' interpretation resulting from this implicature can differ in detail, as the circumstances expected to accompany the event may vary. Hence, different kinds of expected circumstance may be assumed to be lacking, depending on the type of situation described. This variation is illustrated below starting with a very general interpretation in (6). Note for Kambaata that only one of its two manner demonstratives can be used with a 'just so' interpretation.

- (6) a. Kambaata [elicited]

Út-u-s	m-á-ma	kaa'll-áno-a-la?	-
thorn-M.NOM-DEF	what-M.ACC-CF	help-3M.IPFV.REL-M.COP2-PRAG1	
Ut-á	hitt-ínta	le'-is-sáa	bagáan
thorn-M.ACC	like.this[1]-F.ACC<N>	grow-CAUS1-3F.IPFV	but
mexx-u=rr-á-a			
single-M.ACC=NMZ4-M.ACC-ADD			
kaa'll-umb-ó-ssa-a			
help-3M.NEG5-M.PRED-3PL.OBJ-M.COP2			

[Speaker A:] 'Then *what* are the thorns good for?' – [Speaker B:] 'They (= roses) grow them *like this* (= just like this, without a reason), (the thorns) are of no use for them.'

- b. Wan

à	bò	kēé
3SG	leave	like.this

'Leave it *like this*.' (= Do not follow up on it; do not do anything special about it.)

Different aspects of the situation are interpreted as involving only the minimal circumstances, depending on what is most relevant for the particular type of event in a given context. In (6a), in the context of a question concerning the use of the thorns, the answer implies that there is no particular use associated with their growth. In (6b), the situation is taken to involve no particular circumstances in very general terms, and 'just so' can be interpreted as referring to very different possible types of follow-up, depending on what is seen as most relevant in the given context (in the context of a conflict, it may suggest that no retaliation should follow; in the context of leaving a room it may describe leaving the door open, etc.).

In (7), the situation is understood to come about without the expected preparatory stage: in (7a), no customary food preparations precede the event of visiting;

in (7b), the arrival of the intruders is described as sudden, unforeseen, with nothing warning the villagers of their approach.

- (7) a. Kambaata [elicited]

Hitt-ínta mar-áammi-ndo, mexx-u=rr-á
like.this[1]-F.ACC<N> go-1SG.IPFV-Q single-M.ACC=NMZ4-M.ACC
qixx-an-s-u'nnáan?
be.ready-PASS-CAUS1-1SG.NEG4

(The daughter asks the mother whether she is about to visit the circumcised boy. The mother answers:) '(Do you expect that) I go like this, without having prepared anything (to take along)?'

- b. Wan

mጀ mጀ é ŋ zጀ à ŋ blጀ kጀ tጀ kጀé
people PL DEF PERF come 3PL PERF come.out 1PL.INCL on like.this
cáá
QUOT

[Saying:] 'These people arrived, they just came out at us **like this** (= suddenly, without warning).'

In (8), where a buying event is concerned, the relevant circumstances are taken to be the price (as suggested directly by Speaker A's question). Hence, the absence of relevant circumstances is interpreted as implying that, contrary to Speaker A's expectation, no payment was involved.

- (8) Kambaata [elicited]

Me'-íin hi'rr-íti-la? – Hitt-ínta
how.much-M.ICP buy<MID>-2F.PFV.CV-PRAG1 like.this[1]-F.ACC<N>
aass-ée-'e
give-3M.PFV-1SG.OBJ

[Speaker A:] 'For how much did you buy it?' – [Speaker B:] 'He gave it to me **like this** (= for free).'

In (9), the discussion centres on an event of arrival: Speaker A asks Speaker B if, upon their arrival, they brought with them a particular object. The absence of relevant circumstances is most naturally interpreted in this context as implying that there was no accompanying bringing event involved, i.e. even though Speaker B did arrive, the additional event of bringing was not there to talk about.

- (9) Kambaata [elicited]

Ber-e-'ée=bii éeb-bi-ndo? –
 yesterday-F.GEN-ASC.F.GEN=NMZ2.M.ACC bring-2F.PFV.CV-Q
 Kám-be, hitt-ínta-be orbá'
 INTJ-PRAG5 like.this[1]-F.ACC<N>-PRAG5 face.difficulties.1SG.PFV.CV
 waall-oommíi. Mánn-unku-s yóo-ba'a
 come-1SG.PFV.REL.VV people-M.NOM<N>-DEF COP1.3-NEG1
 [Speaker A:] 'Have you brought yesterday's thing?' – [Speaker B
 (disappointed):] 'Far from it! I faced difficulties and came back **like this** (=
 empty-handed, I went there in vain). The people (I wanted to meet) were
 not around.'

In (10), the relevant event is that of hosting a bride at her in-laws' place. The circumstances that are taken as salient, in the context of this particular story, is whether she slept by herself. The use of the manner demonstrative, without any suggested special manner, implies that no special circumstances could be mentioned, i.e. the bride was left in a separate room with no one accompanying her.

- (10) Wan

bé trē gó bō à mō, àjì lē yí tè á kū
 then night inside arrived 3SG PRT 3PL woman sleep kill ADJ.FOC house
 é pāñ dō lé kēé
 DEF side one at like.this

'And when the night came, they put the woman to sleep in a room in the house **like this** (= all by herself).'

Finally, in (11), the situation of yams staying in granaries without any special circumstances is interpreted, in context, as a situation where the yams did not receive proper treatment, and became spoiled.

- (11) Wan

bé à pā yrē ló yā pí wà gē bō
 then 3PL could work do PPS still not PRT then yams PL DEF stayed
 gōñ wā kēé
 granary under like.this

'They could no longer work, and the yams stayed in the granaries **like this** (= rotting away).'

The selection of the examples above illustrates the flexibility of the ‘just so’ interpretation when it comes to determining which aspect of the situation is described as reduced in accompanying circumstances. It is not possible to assign any specific meaning to the manner demonstrative that would fit all such contextual uses. Instead, we believe that the particular interpretation is derived for each use according to the type of event described and the larger context. The mechanism by which the specific meaning is derived is grounded in Gricean principles of relevance and quantity: if the speaker refers to the manner in which the event was realised without suggesting that any special circumstances were involved, they most likely imply that the expected circumstances were absent or significantly reduced.

We leave aside the question of more precise characterisation of the nature of the ‘just so’ interpretation. An approach that seems promising involves treating manner demonstratives as functions evoking contextually salient alternatives in the domain of manner, along the lines suggested by Eckardt (2001) for German intensifying *selbst*. In such an approach, the ‘just so’ interpretation could be related to the contextual choice of the most salient aspects of the situation, roughly corresponding to expected answers to the question “How instead (did it happen)?”. Just as the answer to this question may vary depending on situation type, the ‘just so’ interpretation evokes different kinds of circumstances normally associated with the event.⁴

This line of reasoning could perhaps explain the otherwise puzzling relationship of the Kambaata ‘just so’ demonstrative use with focus: the ‘just so’ reading involves a focus-related *n*-morpheme (glossed *N* in the examples above). On the other hand, the alternative-based approach would run into difficulties explaining the rather specific meaning manner demonstratives receive in the particular examples: rather than implying any kind of surprising manner, as one would expect on the focus-based account, they suggest consistently the *absence* of any particular manner or circumstances with respect to a contextually salient aspect of the situation. This specific reading suggests that the ‘just so’ interpretation may be closer to the *assistive* reading of German *selbst*: ‘by oneself, rather than with the help of others’. This reading does not fit very well with its focus and intensifying meanings, causing Eckardt to introduce a special ASSIST-function for it: the assistive version of *selbst* expresses “the *absence* of any person that stands in the ASSIST-relation to the event in question” (2001: 402). We believe that manner demonstratives show a similar interpretation in the domain of manner: as described above in very general Gricean terms, they signal the absence

⁴We are grateful to Carla Umbach for suggesting to us the relevance of Eckardt’s account.

of any contextually relevant circumstances accompanying the event in question. We leave the formalisation of that interpretation to future research.

3 Differences in extended uses explained by differences in syntax

3.1 Adverbial vs. clause-final marker

Our main goal in this study is to argue that the ways in which new uses of a marker develop from its core use are intimately related to the marker's syntax. As syntactic behaviour determines the word's collocational potential, even minor differences in syntactic category may lead to drastically different paths of development of semantically similar or identical markers. The differences in the use of manner demonstratives of Kambaata and Wan illustrate just that correlation. We would like to argue that the most important difference underlying the markers' different range of uses is syntactic. In Kambaata, manner demonstratives behave as adverbials, while in Wan, the manner demonstrative is a clause-final particle.

As shown in §2, Kambaata has two manner demonstratives, *hittíta* and *hit-tigúta*, glossed 'like.this[1]' and 'like.this[2]', respectively. In their core uses, they are placed in the pre-verbal position (see the examples above), usually immediately preceding the verb. In the corpus data there are a few examples in which they are separated from the verb by other adverbials (12a). They can also be used predicatively (12b), and they appear in all sorts of finite and non-finite constructions, including subordinate clauses with nominalisations, cf. its use in a relative clause in (12c), in a converb clause in (4a), and in adverbial clauses headed by *jeechchúta* 'time' in (5a). All these properties also characterise adverbs.

- (12) a. Kambaata [written]
- | | | |
|--------------------|-------------|------------------------------|
| <i>Hittig-úta</i> | <i>m-íi</i> | <i>xawaaqq-itáyyoont, ko</i> |
| like.this[2]-F.ACC | what-M.DAT | speaker-2F.PROG 2SG.VOC |
| <i>mán-ch-o?</i> | | |
| people-SGV-M.OBL | | |
- 'Hey (my) man, why do you say **this** (lit. like this)?'

- b. Kambaata [oral]
- | | |
|--------------------------|----------------|
| <i>... hanaqal-í=g-u</i> | <i>ikkodáa</i> |
| dish.sp-M.GEN=way-M.NOM | however |

- hittíg-u-ta-ba’á**
like.this[2]-F.PRED-F.COP2-NEG1
‘... but the preparation (lit. way) of the *hanqalú*-dish is not like this.’
- c. Kambaata [elicited]
- Hitt-ítá xe’-áyyoo burtukaan-á
like.this[1]-F.ACC taste.good-3M.PROG.REL orange-M.ACC
xuud-im-bá’á
see-1SG.NIPFV-NEG1
‘I have never seen oranges that taste **this** good (lit. good like this).’

In Wan, the manner demonstrative is not morphologically transparent, and it behaves in a way that is notably different from the adverbial demonstratives of Kambaata. It can only occur in one position within the clause: in the clause-final position following all adverbials, but preceding clause-level particles, as in (13).⁵ Unlike adverbials, it cannot be fronted, and it cannot be used predicatively.

- (13) Wan
- ké mōñ̄ gāā klā tājjí kēé gē ō, mōñ̄ zōñ̄
if LOG.PL went+3SG behind always like.this PRT PRT LOG.PL PROSP
di-ñ̄ à mì wá
reach-PROSP 3SG at not
‘If we always go after him like that, we are not going to reach him.’

The manner demonstrative only occurs in Wan in finite clauses; it cannot occur with nominalisations. Characteristically, the same restriction applies to the use of the negation marker, which is also limited to the position at the end of a finite clause (Nikitina 2009: 923–924). It is explained by the fact that both negation markers and manner demonstratives attach at the IP-level in the clause structure (Nikitina 2008, 2019b). Due to that structural peculiarity, they can only appear in finite clauses, and not with nominalisations.

The manner demonstratives in Kambaata and in Wan differ in one more important respect. In Kambaata, manner demonstratives can occur by themselves, for example in answer fragments in a dialogue. Example (14a) illustrates the use of ‘it is like this’, a non-verbal predicate, as an affirmative answer synonymous to *āā* ‘yes’. In (14b), speaker B gives an answer fragment to A’s first question. No such use is available in Wan: while adverbials commonly occur on their own in similar contexts, the manner demonstrative cannot appear by itself. This difference

⁵Like adverbials, it precedes the negation marker if there is one.

is consistent with the difference in syntactic category: unlike in Kambaata, the manner demonstrative in Wan is not an adverbial but a non-projecting particle, and it cannot occur on its own like lexical constituents.

- (14) a. Kambaata [written]

Mat-ú mat-íin usu'rr-ám-u-a? –
 one-M.ACC one-M.ICP tie<MID>-PASS-M.PRED-M.COP2
 Hittíg-u-ta
 like.this[2]-F.PRED-F.COP2

Speaker A: '(Does it mean) to tie one to the other?' – Speaker B: 'Yes (lit. it is like this).'

- b. Kambaata [elicited]

Káan m-íi hír-teent? M-á
 P_DEM1.M.ACC what-M.DAT buy-2SG.PRF what-M.ACC
 ass-áno-he? – Hitt-ínta
 do-3M.IPFV-2SG.OBJ like.this[1]-F.ACC<N>

[Speaker A:] 'Why did you buy this, what's the use?' – [Speaker B:] 'I bought it just **like this** (= without considering its use).'

All in all, the unusual behavior of the manner demonstrative in Wan suggests that it is not an adverbial, as in Kambaata, but a non-projecting clause-final particle that shares a number of syntactic properties with the negation marker. Table 2 summarises the different syntactic properties of the manner demonstratives of Kambaata and Wan. In the next sections we relate the syntactic difference to differences in the demonstratives' extended use, arguing that as the syntactic category defines the contexts where the marker appears, it ultimately determines the new meanings the marker develops over time.

3.2 Extended uses in Kambaata

There are three types of extended use of manner demonstratives in Kambaata that are not attested in Wan. All three depend on the adverbial status of the relevant markers.

First, the manner demonstrative in Kambaata shows a characteristic interactional use: in the context of a dialogue, one of the two demonstratives, *hittigúta* (like.this[2]), can be used on its own as an affirmative answer ('yes') to the interlocutor's question, as has been demonstrated in (14a). Since the manner demonstrative is an adverbial, it can naturally occur in constructions such as 'It is like this', where the manner demonstrative refers anaphorically to the state of

Table 2: Differences in the syntactic behavior of manner demonstratives of Kambaata and Wan

	Adverbial (Kambaata)	Clause-final particle (Wan)
Position in the clause	Within the verb phrase, preceding the verb	Clause-final only
Finiteness restrictions	Appears in finite or non-finite clauses	Appears in finite clauses only
Independent use	Yes, e.g. as an answer fragment	No

affairs described in the previous utterance. In its interactional use it is marked as a non-verbal predicate. This use results in a conventionalised affirmative form ‘like this’, which may in time become a major interaction-structuring device.⁶

As mentioned in §3.1, no such use is attested in Wan, due to the manner demonstrative’s different syntactic status. Unlike adverbials, particles cannot be used predicatively in Wan. Neither do they project their own constituents that could be used by themselves as answer fragments. The difference from adverbials predicts that the interactional affirmative use could not develop in Wan.

Secondly, the form *hittigúta* (like.this[2]) appears in Kambaata in coordination-like structures, with the meaning ‘also, too, likewise’. In that use, it combines obligatorily with the focus marker *-n*. The coordinated constituents can be sentences or noun phrases (of any syntactic function); see (15a) and (15b).

- (15) a. Kambaata [written]

Hittig-ún̥ta esáa mät-it fít-íchch-ut
 like.this[2]-F.ACC<N> 1SG.DAT one-F.NOM flowering.plant-SGV-F.NOM
 hé’-ee-’e ikkeeráan ...
 exist-3M.PRF-1SG.OBJ IRR

(Preceding sentence: If I had, for instance, a silk scarf, I could wear it around my neck and walk around.) ‘And (lit. like this) if I had a flowering plant, ...’

⁶A reviewer pointed out that this use is very frequent cross-linguistically, citing examples such as Finnish *niin* ‘thus’, Hungarian *igen* (< *íg* ‘so’) and Eastern Slavic, e.g. Polish and Ukrainian *tak* ‘so’. A brief discussion of this use can be found in König (2015).

b. Kambaata [written]

... báar-uhu-u shambalál-uhu-u **hittig-únta**
 sea-M.NOM-ADD lake-M.NOM-ADD like.this[2]-F.ACC<N>
 wól-u uull-á al-í wó'-u
 other-M.NOM earth-F.GEN top-M.GEN water-M.NOM
 hór-unku-u móoll goof-áno
 all-M.NOM<N>-ADD dry.up.3M.PFV.CV finish-3M.IPFV
 ‘... the sea, the lakes **and** (also) other (bodies of) water on the earth
 would all dry up.’

We believe that the coordination-like uses derive from constructions with ellipsis. In the case of intersentential ‘and, also’ (15a), one could assume that a converb clause with a focused proposition-anaphoric manner demonstrative, lit. ‘(it) being (exactly) like this’, was reduced to the extent that only the sentence-initial demonstrative *hittigúnta* was left. In the case of the NP-coordinating ‘so’ (15b), the development in Kambaata is likely to have proceeded as outlined for English *also* in König (2015: 44f). The manner demonstrative likely referred anaphorically to a preceding verbal or sentential constituent (see the manner demonstrative element *so* in *also*), and the two parallel juxtaposed sentences were later simplified through the ellipsis of material that was shared, i.e. ‘The sea would dry up. In the same way other bodies of water would dry up.’ > ‘The sea in the same way (> also, and) other bodies of water would dry up’. The function of intersentential ‘and’ probably preceded the NP-coordinating ‘and’.

Thirdly, the combination of the manner demonstrative *hittítá* (or rather its reduced form *hitt*) and the 3M perfective converb form *íkk* of the verb *ih-* ‘become’ has fused and been lexicalised as an intersentential contrastive connective, *hitt=íkk* ‘but’ (lit. ‘having become so’). This use is illustrated in (16).

(16) Kambaata [elicited]

Hítt=íkk kasal-á-s
 like.this[1]=become.3M.PFV.CV charcoal-M.ACC-3M.POSS
 hi'rr-í=ké' isí kuntaal-iichch ii kuntaal-áan
 buy<MID>-1SG.PFV.CV=SEQ 3M.GEN sack-M.ABL 1SG.POSS sack-M.LOC
 wor-aammí j-áata kasal-i-sí mereer-ón
 put.in-1SG.IPFV.REL time-F.ACC charchoal-M.GEN-3M.POSS interior-M.LOC
 ább-at ább-at háqq-it yoo'i-i xuujj-óomm
 big-F.NOM big-F.NOM wood-F.NOM COP1.3.REL-NMZ1.M.ACC see-1SG.PFV
 (Context: Speaker bought charcoal from a seller who claimed to have the
 best quality product on the market.) ‘But when I had bought his charcoal

and started moving it from his sack into mine, I found out that there were big pieces of wood (hidden) in his charcoal.'

The lexicalisation happened in a construction involving a manner demonstrative in the function of an oblique argument of the verb 'become', lit. 'having become so', where *hitt* 'so' is a propositional anaphor. The frequent use of *hitt=ikk* between sentences that were semantically contrastive caused the contrast to become associated with the element *hitt=ikk* itself. The adverbial's argument status was essential for boosting the frequency of the collocation to the point where it has become stored in the mental lexicon as a separate lexical item, and started a life of its own independent of the manner demonstrative.

In contrast to Kambaata, the manner demonstrative has not undergone lexicalisation in Wan. The absence of parallel uses follows in Wan from the fact that the manner demonstrative is not an adverbial, and cannot function as an argument of any particular verb. As a clause-final marker, it does not form a constituent with the preceding verb and is not closely associated with any specific verbs to the extent that would allow them to undergo lexicalisation typical of collocations.

Table 3 summarises the extended uses attested in Kambaata, along with the suggested path of their development from one of the core uses.

Table 3: Extended uses of manner demonstratives in Kambaata

	Meaning	Path of development
Contrastive connective	'but'	Lexicalisation in a high-frequency collocation, fusion with 'become'
Interactional affirmative	'yes'	Conventionalisation of the anaphoric predicative use '(it is) like this'
Coordination-like use	'likewise, also'	Conventionalisation of the anaphoric predicative with verb ellipsis: '(V) like this NP' = 'same with NP' = 'likewise NP'

3.3 Extended uses in Wan

The manner demonstrative particle in Wan also has uses that are not attested with the manner demonstrative adverbials in Kambaata. We would like to suggest that all these uses can be subsumed under the same type, and follow from the

marker's special syntactic status. In particular, all the extended uses attested in Wan can be viewed as instantiations of the same construction: they involve the manner demonstrative particle introducing a clause-external constituent.

Several types of constituents are attested in this position. First, the manner demonstrative commonly introduces ideophones – words of a special morphosyntactic class that are not integrated in Wan in the clause structure but appear in a clause-external position. Ideophones differ in this respect from other word classes such as adverbs: while adverbs can occur clause-initially or, sometimes, before the verb, ideophones are restricted to the clause-final position; while adverbs appear before the clause-final markers such as the negation marker or the exclamative particle (13), ideophones appear after them, suggesting once again that they are attached to the clause at a higher point in the clause structure.

The examples (17a)–(17b) show how the manner demonstrative refers cataphorically to a manner suggested by an ideophone in a clause external position. This use is not attested in Kambaata, presumably for two reasons. First, as already discussed, the manner demonstrative in Kambaata is an adverbial; adverbials appear in Kambaata before the verb and cannot serve to introduce dislocated or clause-external constituents. Second, ideophones in Kambaata must be introduced by a verb, such as ‘say’ or ‘do’, hence they do not appear in the same syntactic position as in Wan, and they cannot be considered the same syntactic category as the ideophones in Wan, despite their common descriptive denomination.

(17) a. Wan

- klègè kpái é àà glògò á kēé kprò-kpròò-kprò
 chili.pepper real DEF 3PL+3SG crush HAB like.this IDPH
 ‘Real chili pepper, they crush it **like this**: kro-kro-kro.’

b. Wan

- ké è tá-í yī é gó kēé dùùù!
 if 3SG put-HAB water DEF in like.this IDPH
 ‘When he throws himself in the water **like this**: duuu!’

Second, the manner demonstrative particle is used to introduce elements extraposed from the clause. These are most typically numerals and adjectives.⁷ When a numeral appears within the noun phrase, it normally follows the head (18a).

⁷It is difficult to say exactly why this type of extraposition is so prominent in Wan; it could be related to the fact that for historical reasons (Nikitina 2011), Mande languages have a highly reduced VP structure (e.g. there are no ditransitive verbs, as only one argument can be accommodated inside the verb phrase). Heavy constituents tend to appear in extraclausal positions, with co-referring pronouns replacing them in the canonical argument positions.

In (18b) and (18c), it is extraposed from the clause, and a manner demonstrative refers to it cataphorically from within the clause. Wan has no other pronominal that could refer to numerosity, and hence the manner demonstrative seems to be the element that serves best the purpose of substituting a numeral within the clause.

- (18) a. Wan
 gbòlò á
 maiden three
 'three maidens'
- b. Wan
 lāā gbòlò kēé á bē à á gá lé ój
 2SG+ALN maiden like.this three then 3PL COP go PROG firewood
 gá
 search
 'There were once three maidens (lit. maidens **like this**: three), and they were going to search for firewood.'
- c. Wan
 è bē wò kū sòòò à dèjì à ìj bò kēé ój
 3SG that do until IDPH 3SG remaining 3PL PERF stay like.this piece
 dō kpó
 one single
 'And he kept doing it until just one of them was left (lit. until they remained **like this**: a single one).'

Adjectives, like numerals, normally follow the head, and they also have no corresponding pronoun. When extraposed, they are referred to within the clause by the clause-final manner demonstrative (19).

- (19) a. Wan
 nē é, è bōō lé è ìj bàò kēé yèjìglèè
 child DEF 3SG stay+3SG at 3SG PERF become.thin like.this skinny
 'The child, sometime later, grew very skinny (lit. **like this**: skinny).'
- b. Wan
 kē yī é dèjì nì bō á bōjì mì kēé sənì
 if water DEF remaining little stayed RSLT bank at like.this little

kēé ... è gé ...

like.this 3SG said

'When there remained just a little water by the bank like this, little like this ... he said ...'

Structures with extraposition are widely used in narrative discourse to introduce different kinds of what Labov & Waletzky (1967) describe as "evaluation" (see also Labov 1972). Traditional narratives in particular are characterised by such uses, which do not normally appear in elicitation. A rare story consists of narration proper; normally, the storyline is interrupted by the narrator's remarks, explanations, and other types of evaluation. The manner demonstrative often marks such expressions of subjectivity, helping the narrator bridge the gap between the reported events and the current speech situation. In (20a) and (20b), for example, the storyline is related to the moment of narration through the use of a structure with an extraposed element, introduced by a manner demonstrative; such structures are very common in our data.

(20) a. Wan

àñj gà á tròò klà-ñj lāpéá yè péé Bōwlinò tā gōñj
3PL go ADJ.FOC earth pour-NMLZ as.if here and B. on like
kēé
like.this

'[As they were digging the hole,] they went and took out earth like
(lit. like this:) from here up to Bowlino.'

b. Wan

à bò é mì yè lāpéá blè é gōñj kēé bé à
3SG arrive NMLZ at here as.if baobab DEF like like.this then 3SG
klāzí é gé ...
little.sister DEF say

'[When he is coming to trap the woman], he arrives at a distance like
(lit. like this:) from here to the baobab, and her sister says ...'

Table 4 summarises the extended uses attested in Wan that do not occur in Kambaata. All of them can be viewed as instances of the same construction – a structure with right dislocation, where the clause-external element is referred to within the clause by a clause-final manner demonstrative particle. We believe that a likely source for the development of this construction was the discourse-cataphoric use of manner demonstratives illustrated in (3b), where the demon-

strative appears at the end of the clause, referring to an element that follows (as in examples with reported speech).

Table 4: Extended uses of the manner demonstrative in Wan

	Meaning	Path of development
Introducing ideophones		
Introducing displaced numerals		
Introducing displaced adjectives		
Introducing complex adverbial expressions such as comparison expressions	Desemanticised	Discourse cataphoric use, cf. (3)

4 Conclusion

Mechanisms of diachronic change and grammaticalisation paths have been recently gaining importance in different approaches to cross-linguistic variation. Yet theories of grammaticalisation and studies in semantic typology rarely pay close attention to syntax. As we tried to show using the example of manner demonstratives, syntactic aspects of an expression's use can restrict new functions the expressions will develop over time. When it comes to explaining differences in the extent of uses of a semantically identical element in two or more different languages, syntactic category and collocational potential sometimes provide crucial pieces to the puzzle. The difference in the syntactic status of the manner demonstrative in Kambaata (adverbial) and in Wan (clause-final particle) explain why the expressions have strikingly similar core uses but at the same time differ no less strikingly in the range of their extended functions.

Figure 1 summarises our account by means of an *amphichronic* semantic map model that aims at predicting the paths along which manner demonstratives can develop (Jurafsky 1996; Nikitina 2019a). Our fragment of the map only takes into account data from two languages; adding more data will undoubtedly result in a larger and more detailed network of senses. Our contribution to the semantic map approach is the addition of constructional information to the otherwise meaning-based network. While canonical semantic maps focus on meaning and

describe relations between senses in terms of semantic operations (Lakoff 1987, Janda 1990; Jurafsky 1996; Lewandowska-Tomaszczyk 2007; Georgakopoulos & Polis 2018), we believe that information regarding the expression's collocational potential is crucial to describing many types of semantic change. We hope that our perhaps naïve and straightforward way of integrating that information into the semantic map model is but a first step in that direction.

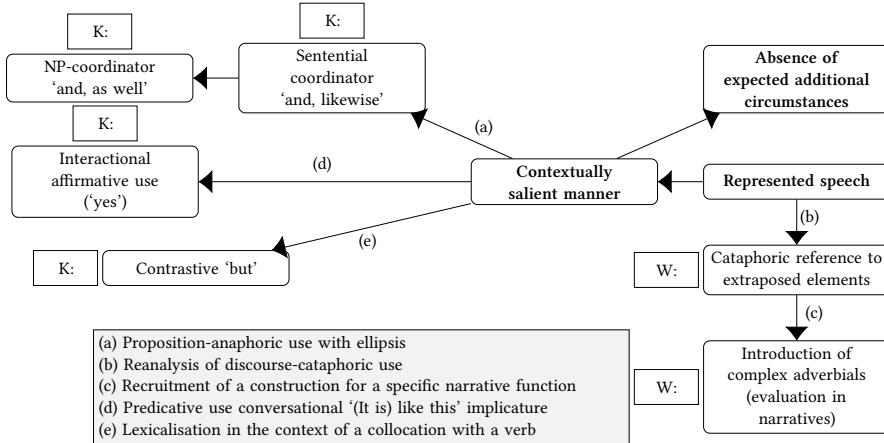


Figure 1: A semantic map enriched by constructional information

As syntactic differences determine differences in collocational potential, they are ultimately responsible for the diverse functions the same expression can become associated with, both at the lexical or morphosyntactic levels and at the discourse-structure level. Our study discussed examples of all these types of change. The adverbial status of the manner demonstrative in Kambaata is responsible for its frequent use as an argument of the verb 'become', causing its combination with the verb to develop into a new lexical expression, the contrastive discourse marker 'but'. The clause-final position of the manner demonstrative in Wan has enabled it to develop into a syntactic device for introducing extraposed constituents of categories that have no other pronominal equivalents (numerals, adjectives, complex adverbials). At the discourse level, the predicative use of the manner demonstrative in Kambaata has given rise to an interactional affirmative use. In Wan, the extraposition-introducing function of the manner demonstrative has enabled it to introduce complex evaluation-related expressions within traditional narrative discourse. All these uses ultimately depend on the demonstrative's distributional potential.

Hence, our study has both methodological and theoretical implications. On the methodological side, it shows that comparison of expressions across languages has a lot to gain from paying close attention to the expression's syntax. When it comes to theory, it serves as a reminder of the need to incorporate fine-grained syntactic information into our models of semantic change. While we tried to make a step in that direction by integrating some constructional information into our semantic map model, further advances should rely on a comprehensive theory of the relationship between syntax and meaning that still needs to be built.

Acknowledgments

Affiliation of the authors: COMUE Sorbonne Paris Cité, INALCO CNRS UMR 8135 LLACAN Langage, langues et cultures d'Afrique. Research for this chapter has been supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 758232).

Abbreviations

Wan glosses

ADJ.FOC	adjunct focus	PL	plural
ALN	alienable possessor	PPS	complement-introducing
COP	copula		postposition
DEF	definite marker	PROG	progressive
DIMIN	diminutive	PROSP	prospective
HAB	habitual	PRT	particle
IDPH	ideophone	QUOT	quotative particle
INCL	inclusive	REFL	reflexive
INTJ	interjection	RSLT	resultative
LOG	logophoric	SG	singular
NMLZ	nominaliser		
PERF	perfect		

Kambaata glosses

A_	adjectival	NEG4	converb negator
ABL	ablative	NEG5	relative negator
ACC	accusative	NMZ1	nominaliser -V
ADD	additive	NMZ2	nominaliser = <i>bii</i>
APPR	apprehensive	NMZ4	nominaliser = <i>r</i>
ASC	associative	NOM	nominative
CAUS1	simple causative	OBJ	object
CF	contrastive focus	OBL	oblique
COND	conditional	ORD	ordinal
COP1	existential <i>yoo</i> -copula	P_	pronominal
COP2	ascriptive/identificational - <i>ha/-ta</i> -copula	PASS	passive
		PFV	perfective
CV	converb	PL	plural
DAT	dative	POSS	possessive
DEF	definite	PRAG1	mitigator <i>-la</i>
DEM	demonstrative	PRAG5	pragmatically
DS	different subject		determined suffix <i>-be</i>
F	feminine		(function as yet
GEN	genitive		undetermined)
ICP	instrumental, comitative, perative	PRED	predicative
INTJ	interjection	PRF	perfect
IPFV	imperfective	PROG	progressive
IRR	irrealis	Q	question
M	masculine	REL	relative
MID	middle	SEQ	sequential
N	focus-related morpheme	SG	singular
NEG1	standard negator	SGV	singulative
		VV	vowel lengthening

References

- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Eckardt, Regine. 2001. Reanalysing *selbst*. *Natural Language Semantics* 9. 371–412.
- Georgakopoulos, Thanasis & Stéphane Polis. 2018. The semantic map model:
State of the art and future avenues for linguistic research. *Linguistics and Language Compass* 12(2). 1–33.

- Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and semantics*, vol. 3: Speech acts, 43–58. New York, NY: Academic Press.
- Guérin, Valérie. 2015. Demonstrative verbs: A typology of verbal manner deixis. *Linguistic Typology* 19(2). 141–199.
- Janda, Laura. 1990. A radial network of a grammatical category: Its genesis and dynamic structure. *Cognitive Linguistics* 1. 269–288.
- Jurafsky, Dan. 1996. Universal tendencies in the semantics of the diminutive. *Language* 72(3). 533–578.
- Karsenberg, Lena & Karen Lahousse. 2018. On the different interpretations of sentence-initial *ainsi* ‘so’ and the competition between three types of verb-subject order. *Folia Linguistica* 52(1). 1–38.
- Keevallik, Leelo. 2005. The deictic *nii* ‘so, in this way’ in interaction. In Daniele Monticelli, Renate Pajusalu & Anu Treikelder (eds.), *From utterance to uttering and vice versa: Multidisciplinary views on deixis*, 109–126. Tartu: Tartu Ülikooli Kirjastus.
- Keevallik, Leelo. 2010. Pro-adverbs of manner as markers of activity transition. *Studies in Language* 34(2). 350–381.
- König, Ekkehard. 2015. Manner deixis as source of grammatical markers in Indo-European languages. In Carlotta Viti (ed.), *Perspectives on historical syntax*, 33–60. Amsterdam: Benjamins.
- König, Ekkehard. 2017. The deictic identification of similarity. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 143–164. Amsterdam: Benjamins.
- König, Ekkehard & Yoko Nishina. 2015. Deixis der Art und Weise, der Qualität und des Grades im Deutschen und Japanischen: Eine kontrastiv vergleichende Analyse. In Yoko Nishina (ed.), *Sprachwissenschaft des Japanischen* (Linguistische Berichte Sonderheft, 20.), 7–31. Hamburg: Buske.
- König, Ekkehard & Carla Umbach. 2018. Demonstratives of manner, of quality and of degree: A neglected subclass. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives: Syntax, semantics and pragmatics*, 285–328. Berlin: De Gruyter Mouton.
- Labov, William. 1972. The transformation of experience in narrative syntax. In William Labov (ed.), *Language in the inner city*, 345–396. Philadelphia: University of Philadelphia Press.
- Labov, William & Joshua Waletzky. 1967. Narrative analysis: Oral versions of personal experience. In *Essays on the verbal and visual arts: Proceedings of the 1996*

3 The use of manner demonstratives in discourse

- Annual Spring Meeting of the American Ethnological Society*, 12–44. Seattle, WA: University of Washington Press.
- Lakoff, George. 1987. *Women, fire, and dangerous things*. Chicago, IL: University of Chicago Press.
- Lewandowska-Tomaszczyk, Barbara. 2007. Polysemy, prototypes, and radial categories. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 139–169. Oxford: Oxford University Press.
- Nikitina, Tatiana. 2008. *The mixing of syntactic properties and language change*. Stanford, CA: Stanford University. (PhD dissertation).
- Nikitina, Tatiana. 2009. The syntax of postpositional phrases in Wan, an “SOVX” language. *Studies in Language* 33(4). 910–933.
- Nikitina, Tatiana. 2011. Categorial reanalysis and the origin of the SOVX word order in Mande. *Journal of African Languages and Linguistics* 32(2). 251–273.
- Nikitina, Tatiana. 2019a. Diminutives derived from terms for children: Comparative evidence from Southeastern Mande. *Linguistics* 57(1). 1–28.
- Nikitina, Tatiana. 2019b. Verb phrase external arguments in Mande: New evidence for obligatory extraposition. *Natural Language and Linguistic Theory* 37(2). 693–734.
- Shor, Leon. 2018. Discourse-anadeictic uses of manner demonstratives: A view from spoken Israeli Hebrew. *Folia Linguistica* 52(2). 383–413.
- Treis, Yvonne. 2019. Similative and equative demonstratives in Kambaata. *Faits de Langues* 51(1). 75–201.

Chapter 4

Morphosyntactic and functional asymmetries in Vatlongos discourse demonstratives

Eleanor Ridge

Massey University

Vatlongos (also known as Southeast Ambrym, Oceanic, Vanuatu) has four demonstrative categories: three person-based distance distinctions (first-person proximal, second-person proximal, and distal) and a contrastive category. In spatial situational domains, and to refer to locative referents, first-person proximal, second-person proximal and distal categories are distinguished from each other. Discourse functions are largely structured around an opposition between forms based on the first-person proximal clitic *ak* and the contrastive suffix *-e*. The wider morphosyntactic distribution of the first-person proximal is reflected in its discourse functions, as the unmarked forms for anaphora and recognitional uses. The more restricted contrastive *-e* forms also occur in contexts of negative affect. In the verbal forms this asymmetry is even more striking. The first-person proximal verbal demonstrative *mak* is the general manner demonstrative, occurring both as a main verb and modifying other verbs in serial verb constructions. The verbal form based on contrastive *-e*, *mue*, is only used in hesitation, a specialisation that could arise from the role of contrastive *-e* forms in discourse repair to modify placeholders.

1 Introduction

Vatlongos, also known as Southeast Ambrym, is an Oceanic (Austronesian) language spoken by around 3,000 speakers in Vanuatu. As well as in the southeast region of Ambrym Island, it is spoken by a community who relocated to Mele Maat, a settlement just outside the capital city Port Vila on Efate Island, in the 1950s. It has many features typical of Oceanic languages: SVO constituent order, head-marking, and subject cross-indexing prefixes and object pro-indexing



suffixes.¹ It has relatively complex morphology, including extensive verb-initial consonant mutation and morphological dependencies in non-contiguous serial verb constructions.

The distribution of Vatlongos demonstratives in discourse shows how extension into different functions can interact with the grammatical properties of morphemes involved, and the wider grammatical structures of the language.

A three-way person-based distance distinction between first-person proximal, second-person proximal and distal is used in the spatial situational domain, but is not maintained in the major discourse extensions of the demonstrative system. Instead, a two-way distinction between first-person proximal and contrastive forms is applied in discourse functions. The first-person proximal is the unmarked form for anaphora, while the contrastive is also used to code negative affect, indicating the speaker's dislike or emotional distress about characters, times and places.

This difference in markedness is also evident in the different grammatical possibilities for the two forms. First-person proximal *ak* is an enclitic that can also function as an independent phonological word, and can freely modify noun phrases (including pronouns), or function as a pronoun. It is also lexicalised with the 3SG pronoun *xi* as *xiak*, which can function as a temporal or locative adverb. The contrastive suffix *-e* cannot appear independently, but must attach to a singular or plural pronoun, formally identical quantifiers which are diachronically related to the pronouns, or more rarely to the locative noun stem *ig*. This means that as an adnominal, it occurs only in singular or plural noun phrases, and not in the dual or paucal. Like first-person proximal *ak*, it can also function as a temporal or locative adverb when attached to the 3SG pronoun or singular quantifier.

First-person proximal *ak* is also a component of the demonstrative verb *mak* 'like this', which is very frequent in the corpus and usually occurs as the subsequent verb in a serial verb construction, meaning '(be/do) like this', 'in this manner'. There is also a verbal form *mue*, which probably originates from the same demonstrative paradigm based on contrastive *-e*, although its use is restricted to a verbal hesitation form. Verbal demonstrative *mue* is used as an inflected placeholder when a speaker is searching for a verbal lexeme.

The examples discussed in this chapter come from a ~65,000 word corpus of audio- and video-recorded texts collected during fieldwork on Ambrym and in Mele Maat between 2014 and 2017, as part of a PhD project focusing on verbal con-

¹This means subject prefixes optionally co-occur with a co-referential noun phrase, whereas object suffixes occur in complementary distribution with object noun phrase (see Haspelmath 2013).

structions (Ridge 2019). The token frequencies of different demonstrative forms are based on their distribution in the spontaneous (non-elicited) texts within the corpus, consisting of a ~48,000-word subcorpus. Most of the texts in this subcorpus can be accessed via the Pangloss collection (Ridge 2018b), which displays audio and video time-aligned by pause unit, with a transcription, interlinear glossing, and translations in English and Bislama (the national language of Vanuatu).

A limitation of the corpus is that it is heavily skewed towards monologues and narratives. This means that some functions of demonstratives are under-represented compared to everyday language use, especially spatial situational uses and discourse functions involving confirmation from an interlocutor. Some of these are pointed out below, and examples are taken from field notes of observed language use to supplement the corpus. However, the functions of demonstratives in relation to referent tracking and establishing topics are very clearly exemplified in these narrative texts.

This chapter focuses on the demonstrative forms that occur most frequently in the subcorpus (see Table 1). These five forms are all based on first-person proximal *ak* and contrastive *-e*, but other demonstrative distinctions and paradigms are outlined in §2. The high token frequency of these forms allows a detailed examination of their discourse functions, whereas comments on some of the other demonstrative forms are necessarily more speculative.

Table 1: Most frequent demonstrative forms in the Vatlongos corpus, their token frequency and major grammatical functions

	<i>ak</i> first-person proximal	<i>-e</i> contrastive
Independent	<i>ak</i> (488 tokens) adnominal locative adverbial	-
3SG/SG <i>xi</i>	<i>xiak</i> (251 tokens) temporal adverbial locative adverbial pronominal adnominal (in complex NPs and relative clauses)	<i>xie</i> (290 tokens) adnominal temporal adverbial locative adverbial pronominal
Verb	<i>mak</i> (328 tokens) subsequent verb in SVC main verb in clause	<i>mue</i> (227 tokens) verbal hesitation form

Vatlongos is most closely related to Paamese, the language spoken on the islands located off the southeast coast of Ambrym, Paama and (historically) Lopevi. Paamese has a simpler demonstrative system, making a two-way distinction between a proximal and a distal. In addition to the expected spatio-temporal functions of those categories, the proximal is used for the “current topic of discussion” in the discourse, and the distal is used to refer back to a previously mentioned referent which is not the current topic of discussion (Crowley 1982: 226–229). Paamese also has verbal demonstratives, and they are more transparently compositional than the Vatlongos equivalents. Crowley identifies a verbal root *muko* which obligatorily occurs with one of the two demonstrative clitics.

The closest parallel to Vatlongos contrastive *-e* is found in Daakaka (West Ambrym), where a clitic *e* ‘the other (place/one)’ indicates contrast regardless of deictic distance (von Prince 2015: 175–176, 333). This form can cliticise to an agent pronoun, a local pro-adverb, and a verb meaning ‘like’, parallel to the distribution of Vatlongos *-e*.

§2 outlines the possible distinctions and forms of Vatlongos demonstratives in spatial situational uses, including the forms that are restricted to the spatial paradigm. §3 focuses on the morphosyntactic distribution of the *ak* and *-e* series, while §4 takes a closer look at the range of discourse functions performed by these forms. §5 focuses on the verbal demonstrative forms, linking their functions to the role of the contrastive *-e* series in discourse repair.

2 Distinctions in Vatlongos demonstrative forms

2.1 Introduction

Earlier descriptions of Vatlongos demonstratives only mention a three-way distinction between the forms shown in Table 2. This paradigm is fundamentally spatial, usually marking deictic reference in situational use, but also available for anaphoric reference to places when combined with the locative bound stem *ig*.

Table 2: The spatial paradigm in Vatlongos demonstratives

Marker	Gloss	Example forms
<i>ak</i>	first-person proximal	<i>ak, xiak, igak</i>
<i>xai</i>	second-person proximal	<i>xai, igaxai, iaxai</i>
<i>xor</i>	distal	<i>xor, igoxor, ioxor</i>

Parker's (1970) dictionary lists *ak* 'this, here' and *xiak* 'now, here', suggesting these forms had roughly similar functions to their contemporary usage. He also lists *xai* 'that (relatively near)' and *xor* 'there, that (relatively far)'. Recent field-work has found that these spatial situational distinctions are person-based, with *xai* marking that the referent is close to the addressee, rather than the purely distance-based medial category implied by Parker's definitions.

Another major difference between Parker's description and the description offered here is the addition of the contrastive series of demonstrative forms based on *-e*. In the contrastive paradigm, *-e* forms are contrasted with forms based on *ak*, as shown in Table 3.

Table 3: The contrastive paradigm in Vatlongos demonstratives

Marker	Gloss	Example forms
<i>ak</i>	first-person proximal, non-contrastive	<i>ak, xiak, igak, mak</i>
<i>-e</i>	contrastive, negative affect, hesitation marker	<i>xie, ige, mue</i>

Like the spatial paradigm, this contrastive paradigm can be used for deictic reference in situational use, but additionally it can be used for discourse deixis, anaphora and recognitional use, functions described in §4. The participation of *ak* forms in both the spatial and contrastive paradigm explains the higher frequency and wider extension of *ak* forms, including into temporal domains. The wider functional range of *ak* forms is also mirrored in their morphosyntactic behaviour described in §3.

The rest of this section will discuss less frequent forms that mark a spatial distinction: the independent forms, the forms based on the locative bound stem *ig*, and the *ia* forms which appear to have developed diachronically as phonologically reduced variants of the *ig* forms. Apart from in the first-person proximal series, these forms are all less frequent in the subcorpus than those in Table 1, which are the main focus of this chapter, so the discussion here is relatively brief and speculative, but important to understanding how the discourse functions of the *ak* and *-e* forms fit into the wider Vatlongos demonstrative system.

2.2 Spatial distinctions in independent forms

Although they are listed in Parker's dictionary, today the independent forms of the second-person proximal and distal demonstratives are fairly marginal (Table 4): there are 21 tokens of *xai* in the corpus and 14 of *xor*, and these tokens are

associated with only a handful of older speakers. They seem to index a conservative style of speech, and are especially associated with the distinctive Endu-Vat-longos dialect, spoken in the northernmost village of Southeast Ambrym (Ridge 2018a). First-person proximal *ak* is far more frequent and has spatial, temporal and discursive functions discussed in §4.

Table 4: The independent forms

Series	Independent form	Gloss	Token count
Proximal 1	<i>ak</i>	near speaker	488
Proximal 2	<i>xai</i>	near addressee	21
Distal	<i>xor</i>	distal	14
Contrastive	-		

2.3 Spatial and contrastive distinctions with locative stem *ig*

In the locative domain, all three spatial forms and the contrastive can be suffixed to the locative stem *ig* ‘place’, a bound stem that can only occur with one of the demonstrative forms, as shown in Table 5.

Table 5: Forms based on locative bound stem *ig* ‘place’

Series	Locative form	Gloss	Token Count
Proximal 1	<i>igak</i>	‘here, this place (near speaker)’	221
Proximal 2	<i>igaxai</i>	‘there, that place (near addressee)’	7
Distal	<i>igoxor</i>	‘there, that place (distal)’	38
Contrastive	<i>ige</i>	‘that place (negative affect)’	21

The first-person proximal form *igak* is by far the most frequent with 221 tokens in the corpus, followed by distal *igoxor* (38 tokens), and the fairly infrequent second-person proximal *igaxai* (7 tokens). However, in everyday interactions, the second-person proximal and distal forms were observed to be more frequent than these figures would suggest. The maintenance of a three-way distance-based distinction in only locative forms (and not demonstrative determiners) has been

observed for Neverver (Malekula, Vanuatu, Barbour 2012). It appears that Vatlongos is headed in this direction given the marginal status of independent second-person proximal *xai* and distal *xor*.

Contrastive -e can also be suffixed to the locative stem, but in this form the negative affect meaning of the suffix seems to be more important than the basic contrastive meaning. Of the 21 tokens of *ige* in the corpus, 18 occur in accounts of emotionally distressing episodes, especially in personal experiences of natural disasters, or to describe places where characters are tricked or disrespected in traditional stories. The three remaining tokens are in the concluding utterances of narratives, in formulaic phrases meaning ‘the story finishes there’. This connotes a humble, apologetic tone at the end of a performance, and co-occurs with other hedging strategies like repetition of the adverb *xos* ‘maybe’.

- (1) [20150419e_h01m128_68]²
 ma xos stori navan xos bus ig-e³
 so maybe story 1SG.POSS maybe 3SG.NFUT.finish place-CONTR
 ‘So maybe my story ends there.’

The *ig* forms can function as locative adverbials, as in (1) and (2), or pronominally as the argument of a verb. In (3), *igak* is the object of the verb: *kamuet* ‘find’ is a morphologically transitive verb which must either be followed by an object noun phrase or take an object pro-indexing suffix.

- (2) [20150415a_h02s125_41]
 lata-pat igak
 3PC.NFUT-sleep place.PROX1
 ‘They slept here.’

²Example codes consist of the date of the recording event in YYYYMMDD format (e.g. 20150419), an identifying letter for each session recorded that day (e), an underscore, a letter indexing broad genre (c = conversation, h = history, n = narrative, p = procedural, t = formal speech), a number identifying the recording within the session (e.g. 01), a letter indicating the speaker community (e = Endu village, m = Mele Maat village, s = other villages of Southeast Ambrym), a speaker code (e.g. 128), an underscore, and a number identifying the pause unit within the recording (e.g. 68).

³Orthography aligns with IPA with the following exceptions: voiced stops are prenasalised ⟨b⟩ /^mb/, ⟨d⟩ /ⁿd/, ⟨g⟩ /^ŋg/; ⟨v⟩ can be realised as [v] or [β], the digraph ⟨ng⟩ represents /ŋ/. ⟨j⟩ represents the affricate /dʒ/ which only occurs in loan words from Bislama and English. Capitalisation and punctuation follow practices in English, the language of education for most Vatlongos speakers.

- (3) [20170413a_h01m169_21]
 la-kamuet igak
 3PL.NFUT-find place.PROX1
 ‘They found this place.’

Like the other forms based on first-person proximal *ak*, *igak* is also used for anaphoric functions, as long as the referent is a place. These anaphoric functions explain the much higher token frequency of *igak* compared to the other *ig* forms. In (4), *igak* is used as recognitional placeholder, followed by the place name Lamap (a village on Epi Island). The listener is expected to be familiar with this place, both because it was introduced six pause units earlier in the narrative, and because the history of the founding of Mele Maat is well-known in the community.

- (4) [20150219b_n01m001_45]
 la-be kalavus e igak, Lamap
 3PL.NFUT-NFUT.go.to prison LOC place.PROX1 Lamap
 ‘They went to prison in that place, Lamap.’

2.4 Second-person proximal and distal forms with *ia*

The second-person proximal and distal forms based on a stem *ia* are probably phonologically reduced variants of the locative *ig* forms. Parker (1970: 8) lists *iaksi* and *ioxor* as variants of *igaxai* and *igoxor*. However, in the contemporary corpus they have a different distribution from the locative *ig* equivalents (Table 6).

Table 6: Forms based on stem *ia*

Series	<i>ia</i> form	Gloss	Token count
Proximal 1	–		
Proximal 2	<i>iaksi</i>	‘there, that place, that one (near addressee)’	15
Distal	<i>ioxor</i>	‘there, that place (distal)’	19
Contrastive	–		

Distal *ioxor* has the greatest functional overlap with the locative form *igoxor*: it occurs as a locative adverbial in eight of the 19 tokens in the corpus. The rest of

the occurrences are adnominal, but all occur in complex NPs, and in six of these *ioxor* occurs in a relative clause with an unambiguously locative meaning.

Of the 15 tokens of second-person proximal *iaxai* in the corpus, only one example functions as a locative adverbial. The two pronominal tokens both refer to a person rather than a place, and in both examples function as predicates, rather than arguments of a verb. The 13 other tokens all modify nominals, and predominantly occur with relative clauses and in other complex noun phrases.

The rest of the chapter will focus on the most frequent *ak* and *-e* forms, which are most prominent in discourse functions.

3 Morphosyntactic distribution of *ak* and *-e*

3.1 First-person proximal *ak* forms

The bare form of first-person proximal *ak* and its variant *ok* are mostly used adnominally, and usually attach to the right edge of a noun phrase, but can also function as temporal adverbials. I am describing *ak* as a clitic, because its distribution is syntactically determined, but it can form a single prosodic word with the preceding word (Zwicky & Pullum 1983; Spencer & Luís 2012). A single main stress occurs on *ak*, as final closed syllables take primary stress in Vatlongos words. However, there are examples where *ak* is an independent prosodic word, especially when used adverbially, which could be an argument for analysing it as a particle instead.

In the glosses below I am aligning it as a separate word to reflect its syntactic status, and using the clitic symbol (=) when it appears to form a single prosodic form with the preceding word. However, this is based on auditory impressions and in some cases, transcribers' decisions about where to write word boundaries, rather than acoustic analysis. The frequent combinations of *ak* with the third-person singular pronoun or singular quantifier *xi*, and the locative stem *ig* discussed above, are lexicalised forms glossed as single words.

First-person proximal *ak* follows a noun (5) or pronoun (6) heading a noun phrase.

- (5) [20170331b_n01s034_30]
atou =ak
woman =PROX1
'this woman'

- (6) [20141107b_n01e020_19]
xalu =ak
3DU =PROX1
'these two'

In more complex noun phrases, *ak* can follow a variety of word categories, including adjectives (7), quantifiers (8)–(9), and possessive classifiers (10).

- (7) [20170222f_n01s153_30]
[ses [xa val]] ak
shellfish REL big PROX1
'this big shellfish'
- (8) [20141220g_n01s080_03]
[[horamue nalu] xalu] =ak
boy 3DU.POSS DU =PROX1
'these two boys of theirs'
- (9) [20170413e_n01m030_21]
[[holesok sap] xil] =ak
thing different PL =PROX1
'these different things'
- (10) [20170221e_n01o150_04]
[horamue nan] =ak
boy 3SG.POSS =PROX1
'this boy of his'

Similarly, *ak* can follow any word at the end of a preposition phrase, like the noun in (11), or at the end of a relative clause, like the preposition in (12). This shows that it is syntactic boundaries, rather than word class, that determines the position of *ak*, supporting the analysis as the clitic.

- (11) [20141106f_n01e018_07]
[vul [te oum]] =ak
hole of crab =PROX1
'this crab hole'

- (12) [20170222f_n01s153_49–50]

[venu [xa di gan e nesau]] ak
 volcano REL CONT 3SG.NFUT.burn LOC up PROX1
 ‘this volcano that’s burning up there’

In all the examples so far, *ak* occurs at the right edge of a noun phrase. However, within a noun phrase, it can be followed by a relative clause (13), a coordinator (14), or *mun* ‘too, also’ (15).

- (13) [20150219b_n01m001_44]

[xil =ak] [xa [tahal Maat]]
 3PL =PROX1 REL from Maat
 ‘the ones from Maat’

- (14) [20141106f_n01e018_26]

[[[tut atuli] =ak] xal [mama nan]]
 little girl =PROX1 with mum 3SG.POSS
 ‘this little girl and her mum’

- (15) [20141106f_n01e018_21]

[[[atal =ak] mun] be xalu
 snake =PROX1 too 3SG.NFUT.go.to 3DU
 ‘The snake also went to them.’

Finally, *ak* occurs as a locative adverb modifying a clause, as in (16).

- (16) [20170413a_h01m169_95]

na-taa-ra =ti ak
 1SG.NFUT-NEG-stay =NEG PROX1
 ‘I wasn’t here.’

The lexicalised form *xiak* consists of *ak* and the form *xi*, which is both the third-person singular pronoun and the singular quantifier. This form can occur in many of the same syntactic environments as independent *ak*. It can occur adnominally at the right edge of a noun phrase. However, compared to bare *ak*, *xiak* is more likely to occur when the noun phrase also contains other modifiers (17).

- (17) [20170220e_n01s147_12]

vul venu xiak
hole volcano PROX1
'this volcano crater'

In its adnominal function, *xiak* is more likely to follow the relative clause marker *xa* than to adjoin directly to the noun phrase. This strategy is also preferred for most adjectives in Vatlongos. Unlike in main clauses, adjectives, noun phrases and demonstratives in relative clauses are rarely preceded by the copular verb (18).

- (18) [20141106f_n01e018_15]

mi-gur oum xa xiak
3SG.NFUT-NFUT.take crab REL PROX1
'She took this crab.'

It is also fairly common for *xiak* to occur after a relative clause, referring back to the head of the noun phrase, see (19) and (20). Adnominal demonstratives are often reanalysed as relative clause boundary markers cross-linguistically (Diessel 1999: 132–135).

- (19) [20170222d_n01s152_71]

ueili [xa tati ba te-pou-e e Saot-Is] xiak
pig REL dad GO 3SG.PRI-carry-3OBJ LOC Southeast PROX1
'the pig that Dad went and got from the Southeast'

- (20) [20170315b_n01m159_15]

iat [xa ma-ling joj daen] xiak
yard REL 1PL.NFUT-put church 3SG.NFUT.stay.in PROX1
'the yard that we put the church in'

These tendencies show that while there is overlap in the adnominal functions of *ak* and *xiak*, they tend to occur in slightly different syntactic environments. As might be predicted on the basis of its longer form, *xiak* seems to have greater syntactic prominence, and may help hearers to process syntactic boundaries in complex noun phrases.

Xiak is used as a locative adverbial more frequently than bare *ak*. It can refer to a location that is near a real or fictional place of speech, or to a place that is already activated in the discourse, see (21) and (22). In this use it also has functional overlap with *igak*, discussed in §2.

- (21) [20141117a_n01m003_50]

taa-ve vonine-mem =ti xiak
 3SG.NFUT.NEG-NEG.COP place-1PL.EXCL.POSS =NEG here.PROX1
 'It's not our place here.'

- (22) [20170124b_t01e137_15]

metilou da xiak
 nephew 3SG.NFUT.stay here.PROX1
 'Nephew is here.'

Unlike bare *ak*, *xiak* is also used as a temporal adverbial, see (23) and (24). In the corpus, the locative adverbial uses can be difficult to distinguish from the temporal ones, but *xiak* seems to be most frequently used with a temporal interpretation. In its temporal extensions it can refer to the time of real or fictional speech, or to the time frame that is under discussion in the discourse, which may be in the past, as in (25).

- (23) [20150419e_h01m128_24]

di a-mmei xiak?
 CONT 2SG.NFUT-come now
 'Are you coming now?'

- (24) [20170217j_n01s125_19]

eai ba di mi-nu xiak
 sun GO CONT 3SG.NFUT-dive now
 'The sun is setting now.'

- (25) [20141117a_n01m003_17]

kavmen mi-leh moletin xil tu la-mmei Epi xiak
 government 3SG.NFUT-take person PL already 3PL.NFUT-come Epi now
 'The government had already brought people to Epi then.'

Xiak can additionally be used as a pronominal form that can function as the argument of a verb. In (26), *xiak* is the subject of the verb and refers to a person present in the speech situation (the researcher), who is standing beside the speaker. In (27), *xiak* is the object of the instrumental preposition *ni* and refers to a loan already mentioned in the autobiographical narrative.

- (26) [20141028a_c01m002_07]
xiak bitene va-hur sepin-iен nar
3SG.PROX1 3SG.NFUT.want 3SG.IFUT-take talk-NMLZ 1PL.INCL.POSS
'This one wants to learn our language.'
- (27) [20170413e_n01m030_19]
na-staat-ni pisnis navan ni xiak
1SG.NFUT-start-TR business 1SG.POSS INS 3SG.PROX1
'I started my business with this.'

3.2 Contrastive *-e* forms

I analyse contrastive *-e* as a suffix rather than a clitic because it is heavily restricted in the forms it can attach to, and its syntactic distribution is dependent on the host forms, rather than independently attaching to a syntactic constituent (Zwicky & Pullum 1983; Spencer & Luís 2012). Contrastive *-e* is suffixed to a subset of the pronouns and formally identical quantifiers, as well as the locative stem *ig* discussed in §2.3. Unlike first-person proximal *ak*, syntactic requirements do not ever directly determine the position of *-e*: it can only occur in positions available to the forms it is suffixed to. This is a possible counterexample to Diessel's (1999: 25) hypothesis that bound demonstratives are always clitics rather than affixes, although semantically the suffixed forms do always modify or substitute for a phrase as he argues.

Contrastive *-e* is most frequently suffixed to *xi*, which is the third-person singular pronoun or definite singular quantifier, or *xil*, the third-person plural pronoun or plural quantifier. The singular pronoun form is shown in (28), as the object of the instrumental preposition *ni*. The plural pronoun form is shown in (29), functioning as the subject of the verb.

- (28) [20141116b_c01m_26]
lu-bunit ni xi-e
3DU.NFUT-NFUT.cover.laplap INS 3SG-CONTR
'They covered the laplap⁴ with that.'
- (29) [20141208a_n01m045_29]
xil-e la-pangei nou
3PL-CONTR 3PL.NFUT-spoil 1SG
'They disrespected me.'

⁴Laplap is a food made from grated root vegetables or banana soaked in coconut milk and baked in leaves.

4 Morphosyntactic and functional asymmetries in Vallongos

The singular and plural quantifier forms are shown in (30) and (31), modifying a single noun.

- (30) [20141027a_n01m001_109]

tovolih xi-e
old.woman SG-CONTR
'that old woman'

- (31) [20150310a_h01s114_08]

meseau xil-e
fish PL-CONTR
'those fish'

These forms have the same distribution within the noun phrase as unmarked quantifiers, usually following any adjectives (32) or possessive classifiers (33). However, the contrastive form of the singular quantifier is much more frequent than the singular quantifier alone, which only occurs twice with an unambiguous quantification function. It therefore serves as a host for the contrastive demonstrative suffix, rather than making an independent semantic contribution.

- (32) [20170220g_n01s148_24]

mai mieh xi-e
reef white SG-CONTR
'that white reef'

- (33) [20141027a_n01m001_29]

holesok salu xil-e
thing 3DU.POSS PL-CONTR
'those things of theirs'

Xie is especially common at the end of long relative clauses, and more frequent than *xiak* in this environment (34).

- (34) [20141220g_n01s080_11]

ma-be [[rute [xa xouk o-bit
1DU.EXCL-NFUT.go.to place REL 2SG 2SG-NFUT.say
mal-naa-va =ti en]] xi-e]
1DU.EXCL.FUT-NEG-go =NEG in SG-CONTR
'We went to that place that you told us not to go to.'

Like *xiak*, *xie* is also often introduced by the relative clause marker *xa*, where it is ambiguous between a quantifier and pronominal reading (35).

- (35) [20141107b_n01e020_31]

tatal xa xi-e
snake REL (3)SG-CONTR
'that snake'

Xie can also modify a clause as a temporal (36) or locative (37) adverb, which is evidence of lexicalisation, as in the case of *xiak*. Although it can again be difficult to distinguish between temporal and locative meanings in individual examples in the corpus, the temporal meaning of the adverb appears to be more frequent.

- (36) [20141208a_n01m045_42]

mu-lul vatang xie
3SG.NFUT-shake a.lot then
'It [the earthquake] shook a lot then.'

- (37) [20141219c_n01s033_21]

xale-n daen xie
tail-3SG.POSS 3SG.NFUT.be.in.it there
'Its tail was in there.'

The morphosyntactic distributions of contrastive *xie* and *xile* therefore appear to encompass the various functions of both *ak* and *xiak* in the first-person proximal series, while making an additional distinction for number, but are not able to modify dual and paucal referents. There is a single example in the corpus of *-e* modifying the first-person plural inclusive pronoun *xir*.

Finally, *-e* can modify the locative stem *ig*, as discussed in §2.3.

3.3 Summary

The morphosyntactic distributions of these forms are summarised in Table 7. The greater syntactic independence and freer distribution of first-person proximal *ak* is reflected in the wider discourse functions of this series.

Table 7: Summary of morphosyntactic distributions of *ak* and *-e* forms

Morphosyntactic distributions	<i>ak</i> forms	<i>-e</i> forms
adnominal	<i>ak</i> , <i>xiak</i>	<i>xie</i> , <i>xile</i>
after relative clause marker	<i>xiak</i>	<i>xie</i> , <i>xile</i>
after complex NP	<i>xiak</i>	<i>xie</i> , <i>xile</i>
locative adverbial	<i>ak</i> , <i>xiak</i> , <i>igak</i>	<i>xie</i> , <i>ige</i>
temporal adverbial	<i>xiak</i>	<i>xie</i>
pronominal	<i>xiak</i>	<i>xie</i> , <i>xile</i>

4 Discourse functions of *ak* and *-e*

4.1 Introduction

The wider morphosyntactic distribution of *ak* is mirrored in its use in semantically general, high frequency discourse contexts. *Ak* is the default form for anaphora, with an important role in referent tracking, especially in establishing topics. It is frequent for second mention of new referents, and for switch topics. Contrastive *-e* can occur in some of these discourse environments, but also when an explicit contrast is being made, or in contexts of negative psychological affect, for example if a switch topic is also an adversary in a narrative.

Ak forms are associated with recognitional uses, for referents that the speaker asserts to be known to the hearer, while anticipating that the information given may not be sufficient (Himmelmann 1996: 230; Diessel 1999: 105–109). These are not well-exemplified in the corpus, but were observed during fieldwork.

In this section it is sometimes useful to include a translation of the section of discourse preceding or following an example; this is given in the translation line, in square brackets.

4.2 First-person proximal *ak* forms

If any demonstrative is used in coding a referent, *ak* forms are the default for anaphora. This is especially likely when establishing a topic, or for switch topics when a referent is active in the discourse, but not currently under discussion. Continuing topics are usually coded only by subject-indexing verbal prefixes.

In (38) (which is the fuller context for (27)), *xiak* is first used adnominally to modify the complex noun phrase ‘money that I got’, which refers back to a loan

first mentioned two pause units earlier. Then *xiak* is used again to refer to the money pronominally.

- (38) [20170413e_n01m030_19]
- | | | | |
|--------------------|----------|-----------------------------------|-------|
| pipin-ni | mani | taxa na-gur-i | xiak. |
| 3SG.NFUT.enough-TR | money | REL 1SG.NFUT-NFUT.take-3OBJ PROX1 | |
| na-staat-ni | pisnis | navan ni xiak | |
| 1SG.NFUT-start-TR | business | 1SG.POSS INS 3SG.PROX1 | |
- ‘It was as much as I could buy with the money that I got. I started my business with this.’

Adnominal marking with *ak* forms is also very frequent at second mention of a referent that remains prominent in the subsequent discourse. In (39) (which is the fuller context for (8)), the two boys are introduced in a noun phrase without any demonstrative. Anaphoric *ak* modifies ‘their two boys’ on the second mention, which Lichtenberk (1996: 385) calls “immediate anaphora after first mention”. They are then referred to with the third-person dual pronoun *xalu*.

- (39) [20141220g_n01s080_02-03]
- | | | | | |
|---|---------|----------|--------------|---------------------|
| lu-pes | horamue | nualu | lu-be | lu. |
| 3DU.NFUT-bear | boy | 3DU.POSS | 3DU.NFUT-COP | two |
| lu-pes | horamue | nualu | xalu ak, | lu-laxat |
| 3DU.NFUT-bear | boy | 3DU.POSS | DU PROX1 | 3DU.NFUT-look.after |
| ‘They had two boys. They had these two boys of theirs, they looked after them.’ | | | | 3DU |

This is an example of a tail-head linkage structure (de Vries 2005), described as bridging constructions in a recent typology (Guérin 2019). This is a common strategy for organising discourse in Vanuatu languages (Early 1994: 454; Hyslop 2001: 426–427; Thieberger 2004: 324–325; Schneider 2010: 240; Brotchie 2009: 298–299; Jauncey 2011: 376). A situation or referent that is new information at the end of one section of discourse is repeated as old information at the beginning of the next section of discourse, as background for further new information. In Vatlongos, adnominal *ak* often occurs in the repeated stretch of discourse.

Example (40), from the beginning of an account of tropical cyclone Pam in 2015, shows *ak* modifying the ‘cyclone’ that has been introduced in the preceding clause, in a topicalised clause-initial noun phrase. The topicalised noun phrase precedes the independent pronoun in subject position *xamem* ‘1PL.EXCL’, and is marked prosodically with rising intonation and a pause (indicated by a comma in the transcription line).

- (40) [20150410a_h01s121_02-03]

na-sep ulis =ti vatiang xa bas xir.
 1SG.IFUT-say about =PART cyclone REL 3SG.NFUT.hit 1PL.INCL
 vatiang ok, xamem di ...
 cyclone PROX1 1PL.EXCL CONT

'I'm going to talk a bit about the cyclone that struck us. This cyclone, we were ... [getting news about it].'

The very high frequency of *ak* in these anaphoric uses suggests it could be reanalysed as an article, in opposition to *tei* 'one', which functions like an indefinite article. This is similar to the role of the most general deictic particle in Lewo (Epi, Vanuatu) (Early 1994: 225).

Similarly, *xiak* is the unmarked form for the locative and temporal adverb, and can modify fairly distant situations in space and time. While in examples (21) to (24) *xiak* refers to the here and now of the speech event, it can also refer to times and places under discussion regardless of distance from the speech event, as in (25) where *xiak* refers to a time several decades earlier.

Recognitional use of *ak* was very frequently observed in casual conversation in the field, in contexts like (41), when a speaker is explaining something and wants to make sure that the listener has understood the intended referent. The speaker is also suggesting that the listener does know about the little house in question.

- (41) [field notes]

nim hokkorong tang nesau ak
 house small just up PROX1
 'that little house up there'

However, there are few clear-cut examples of this in the corpus. Their low incidence may be due to the presence of the researcher at most recording events: speakers have low expectations of the researcher-as-listener's cultural knowledge and shared frame of reference, and are more likely to introduce a referent as new information rather than assuming that the whole audience has prior knowledge. Another difficulty in identifying examples of recognitional uses in the corpus is that many possible examples could instead be analysed as anaphora, as the referent has usually been mentioned at some earlier point in a longer narrative, a difficulty observed by Himmelmann (1996: 236). Himmelmann (1996: 235) also observes that distal demonstratives are far more common in recognitional uses cross-linguistically, so the role of *ak* in these uses is suggestive of how general and unmarked the Vatlongos first-person proximal is in its discourse extensions.

This supports the analysis that, in these uses, *ak* is in a paradigmatic relationship with the contrastive, rather than with the other spatial demonstrative series.

First-person proximal *ak* is also associated with what could be described as “coercive recognitional uses”, to assert that the hearer should be familiar with the referent, even if the speaker suspects that they are not. This was very frequently observed in speech directed to me by my host families, especially in relation to kinship networks. Example (42) is a typical example from observation – it was obvious that I did not know which aunt was being referred to, but the use of *ak* made it clear that I should have known.

- (42) [field notes]
tila nam ak
aunt 2SG.POSS PROX1
'this aunt of yours'

The forms based on *ak* and *-e* forms have complementary roles in discourse repair, which are important to the distinction marked in the verbal demonstratives discussed in §5. *Ak* forms often mark a constituent that has been provided after a hesitation. While a speaker is searching for a word, the nominal hesitation form *na* can be prosodically lengthened for as long as it takes to find the target word. When the target word is spoken, it is then often marked with the *ak* form, as in (43).

- (43) [20150305h_h01o111_12]
lu-pus na nim ok
3DU.NFUT-see HES house PROX1
'They saw um, this house.'

4.3 Contrastive *-e* forms

The primary function of the *-e* forms is to mark an explicit or implicit contrast. Often the contrast is with a referent marked with first-person proximal *ak*, but not necessarily. In (44), a father is giving instructions to his sons to hunt on ‘this side’ (marked with *ak*), contrasted with ‘our home’, which is not in the area they should hunt in.

- (44) [20141220g_n01s080_11]
tim saratel xi-e
home 1PC.INCL.POSS SG-CONTR
'[You must go hunt on this side (*ak*) because of] our home there.'

In (28), the contrast is implicit, rather than explicitly stated in the discourse. The pronominal use of *xie* refers to leaves of the *nelnel* plant, as opposed to *hahau* leaves, which are usually used in preparing laplap. Earlier in the conversation, the speaker explained how no one went to collect the *hahau* leaves.

Contrastive -e forms are also used to index negative affect, especially for situations associated with emotional distress or disrespect, and referents that are figured as adversaries in a narrative. The role of negative affect in the use of demonstratives is often figured in terms of psychological distancing. For example, Early (1994: 225) aligns “spatial, temporal and psychological location” in his description of demonstratives in Lewo (Epi, Vanuatu). However, in Vatlongos the demonstrative associated with these functions does not (at least synchronically) have a spatial primary meaning, so the metaphorical rationale for this polysemy is not one of physical distance.

Xie often marks the “enemy” character in traditional narratives, frequently an old woman (30) or a snake (35). It can be used more generally for troublemakers, as in (45), referring to an ancestor who caused a volcanic crater to open by eating forbidden sugarcane.

- (45) [20170220e_n01s147_09]

avu	xa	xi-e	mei	pol-ni	travol
grandparent	REL	(3)SG-CONTR	come	3SG.NFUT.make-TR	trouble
‘That ancestor came and made trouble.’					

Contrastive -e forms are closely linked to the breaking of social conventions, especially taboos and concepts of respect and obedience, as shown in several examples above. In (29), *xile* refers to a group of people – a village on another island – who tricked the protagonist into sleeping with his own mother. In (34), two brothers are explaining to their father that they disobeyed his command, and *xie* modifies the place that they were forbidden to hunt in. In (37), *xie* is used as a locative adverb, and refers to the rat’s bottom, a taboo body part.

As a temporal adverb, *xie* is especially associated with times of fear and distress, as in (36), hardship (46) or discomfort (47).

- (46) [20150219b_n01m001_08]

te-staat	e	fotitu	xie
3SG.PRI-start	LOC	42	then
‘It [World War Two] started in 1942 then.’			

- (47) [20150129c_i01m095_29]
ut mi-ttin ma na-po rat-i xie
place 3SG.NFUT.hot so 1SG.IFUT-carry out-3OBJ then
'It's hot so I'll take it out now.'

Neither the temporal distance of the time indexed by the temporal adverb, nor the relative tense marking of the clause, seems to influence the use of *xie*. In the examples here it indexes a time a few weeks ago (36), many decades ago (46), and just after speech time (47), and modifies verbs marked with the non-future, the prior, and the immediate future, respectively.

However, *xie* as a temporal adverb is especially likely to co-occur with strategies for marking completed situations. This could be an extension of the basic contrastive meaning of the -*e* forms, as an emphasis on the boundaries of events allows them to contrast with other events in a sequence. In (48) and (49), *xie* co-occurs with the adverb *turei* 'already'. Example (49) also includes the use of the verb *hus* 'finish' (here the non-future affirmative form *bus*), which is a high frequency strategy for marking completive aspect in Vatlongos. Example (50) uses the combination of negative polarity and *mu* 'first' to express 'not yet'.

- (48) [20141219c_n01s033_08]
ale lanstaem turei xie, va ra-vanien
so lunchtime already then GO 1PL.INCL.IFUT-eat
'Okay, it's lunchtime already now, let's go eat.'

- (49) [20141212f_n01s054_14]
meu na-pein-ni xouk turei bus xie
megapode 1SG.NFUT-paint-TR 2SG already 3SG.NFUT.finish then
'Megapode I've already finished painting you now.'

- (50) [20150219b_n01m001_84]
lut-naa-ve pasta =ti mu xie
3DU.PRI-NEG-COP pastor =NEG first then
'They weren't pastors yet then.'

Contrastive -*e* forms have a different role to play in discourse repair. Whereas first-person proximal *ak* forms mark a successful repair after a hesitation, -*e* forms modify placeholders which are used instead of more specific formulations,

or because a speaker has decided not to continue searching for an intended lexical item. In (51), *xie* modifies the general nominal placeholder *neta* ‘thing’ (which follows the bound noun *relit* ‘egg’, obligatorily followed by a possessor). At no point in the narrative is this mysterious object, which later becomes the volcano, given a precise name, although it is compared to a chicken’s egg. In a different telling of the same narrative it is instead referred to as a shell.

- (51) [20170220g_n01s148_38]
 na relit neta xa xi-e
 HES egg thing REL (3)SG-CONTR
 ‘[They took out] um that thing’s egg.’

4.4 Summary

Table 8 summarises the discourse functions described in this section. It is striking that *ak* forms have many more discourse functions than contrastive *-e* forms, in addition to their role in the spatial paradigm. This large functional extension suggests a reason for the greater morphosyntactic flexibility of *ak*, which has more functions to perform in a variety of discourse contexts.

This section has shown the complementary roles of these two demonstrative series in the mechanics of discourse repair after a hesitation or false start. First-person proximal *ak* is used for discourse repair, when a referent is repeated after

Table 8: Summary of discourse functions of *ak* and *-e* forms

Discourse functions	<i>ak</i> forms	<i>-e</i> forms
default anaphora	✓	
establishing topics	✓	
second mention	✓	
switch topics	✓	
contrastive topics		✓
repeated content in tail-head linkage	✓	
recognitional uses	✓	
coercive recognitional uses	✓	
successful repair after hesitation	✓	
placeholders for failed repairs		✓
negative affect		✓

a false start. Contrastive *-e* is instead used with placeholders when the speaker is searching for a word. This distinction is important to understanding how the restricted functional distribution of the verbal demonstrative *mue* discussed in §5.4 could have arisen.

5 Demonstrative verbs

5.1 Introduction

In addition to the adnominal, adverbial and pronominal forms described so far, Vatlongos has a verbal manner demonstrative *mak* based on first-person proximal *ak*, and a verbal hesitation form *mue* which could have developed from contrastive *-e*, in line with the contrast between *ak* and *-e* forms involved in discourse repair (Table 9). There is also one token in the corpus of a verb *maxai*, a second-person proximal *xai* form.

Table 9: Verbal demonstrative forms in Vatlongos

Series	Verb	Gloss	Token count
Proximal 1	<i>mak</i>	‘like this’	328
Proximal 2	<i>maxai</i>	‘like that’ (associated with addressee)	1
Distal	-		
Contrastive	<i>mue</i>	hesitation	227

Paamese, the most closely related language, also has verbal demonstratives, which Crowley (1982: 229) analyses as comprising of a deictic verbal root *muko*, and the demonstrative clitics, *=ke* (proximal) and *=neke* (distal). Unlike in Vatlongos, the verbal root can be separated from the demonstrative elements by an intervening negative-partitive suffix and the additive clitic, but *muko* cannot occur without one of these two clitics.

5.2 First-person proximal *mak*

The first-person proximal verbal form *mak* functions as a manner demonstrative. It is very frequent, occurring roughly once every 150 words (328 tokens in the subcorpus). Its high frequency and broad range of functions mirrors the extension of *ak* in identifying discourse referents, but instead refers to manner.

As a main verb it is especially common in procedural texts, referring to a manner being demonstrated non-verbally. In (52), the speaker is demonstrating how to make simboro, grated root vegetables or banana rolled in cabbage leaves and cooked in coconut milk.

- (52) [20141212g_p01s046_05]
 mak, di, mak tang
 3SG.NFUT.like.this then 3SG.NFUT.like.this just
 'Like this, then, just like this.'

Mak is also often used to introduce performances, as in (53) where it first introduces a rope drawing of a nut, and then the song that accompanies the drawing. When a performance is verbal, this can be thought of as discourse deixis, referring to the words of the song. Although in these two examples the direction of reference is cataphoric, *mak* can also refer backwards in time to a completed performance.

- (53) [20141222c_p01s087]
 tavu mak: [...] ale sisien nen mak: [...]
 bushnut 3SG.NFUT.like.this so song of.it 3SG.NFUT.like.this
 'The bushnut goes like this: [rope drawing]. Then the song of it goes like this: [song].'

It is also used for anaphoric discourse deixis referring to situations expressed in the preceding discourse. In (54), *mak* refers to the preceding sentence (26), i.e. that I want to learn Vatlongos, as an explanation for the speaker's actions. The use of discourse deictic *mak* followed by *ma* 'so, then' is a fairly frequent strategy for expressing causal or purposive links between situations in the discourse.

- (54) [20141028a_c01m002_08]
 mak ma na-kes xamim
 3SG.NFUT.like.this so 1SG.NFUT-call 2PL
 'It's like this so I called you.' (i.e. 'That's why I called you.')

This discourse deictic use can also refer to the speech of another interlocutor in a conversation, as in (55). *Mak* collocates with *tang* 'just', often to express a general affirmation, which could be translated as 'that's it' or 'that's the way' in English. This collocation occurs 27 times in the corpus, representing nearly a tenth of tokens of *mak*, and is probably even more frequent in casual conversation, which is underrepresented in the corpus.

- (55) [20150129b_i01m094_15]
- | | | |
|--------------------|---------------|--------------------|
| mu-tin | nahou | tang – |
| 1PL.EXCL.DFUT-burn | garden | just |
| a | bos, | mak |
| oh | 3SG.NFUT.good | 3SG.NFUT.like.this |
| just | | |
- (A:) ‘We’ll just burn the garden.’ – (B:) ‘Oh good, that’s the way.’

However *mak* most often occurs as the subsequent verb in a serial verb construction, modifying the manner of the situation expressed by the initial verb. *Mak* always takes third-person singular subject agreement marking in these constructions, which is the default for event-argument⁵ serialisation in both Vatlongos (Ridge 2019: §6) and Paamese (see Crowley 1987: 61 and Crowley 2002: 61 on ambient core-layer serial verb constructions in Paamese). In Vatlongos serial verb constructions, the relative tense marking on the initial verb determines the marking on the subsequent verb. Matching relative tense marking is required in all the examples of serialised *mak* here, because they are in affirmative polarity.

In serialisation, *mak* can be used exophorically to refer to the external situation, as in (56), from a conversation recorded while eating laplap on a Sunday.

- (56) [20141116b_c01m_54]
- | | | | | | |
|----------|-------|--------|-------------------|--------|--------------------|
| xir, | evri | Sade | ra-ga | laplap | mak |
| 1PL.INCL | every | Sunday | 1PL.INCL.NFUT-eat | laplap | 3SG.NFUT.like.this |
- ‘Us, every Sunday we eat laplap like this.’

Serialised *mak* also refers to non-verbal demonstrations, both in procedural demonstrations (57), and for more spontaneous demonstrations and iconic gestures. In (58), the speaker picks up a banana to demonstrate how a character in a narrative blocks a snake’s mouth with a nut.

- (57) [20141212g_p01s046_04]
- | | |
|--------------------|--------------------|
| la-pis-i | mak |
| 3PL.NFUT-roll-3OBJ | 3SG.NFUT.like.this |
- ‘They roll it like this.’
- (58) [20141220g_n01s080_52]
- | | | | | | | |
|------|----------|---------------|-------|---------|-----|--------------------|
| tati | natel | gur | huit | tavu | tei | mak |
| dad | 3PC.POSS | 3SG.NFUT.take | fruit | bushnut | one | 3SG.NFUT.like.this |
- ‘Their Dad took a bushnut like this.’

⁵I am following Aikhenvald’s (2006: 18–19) typological framework in using this term, rather than “ambient”, which is used by Crowley to describe the same argument structure configuration.

Serialised *mak* can also be used for discourse deixis. In (59), the preceding stretch of discourse describes how the dwarf has been pulling the rat's tail, and *mak* refers back to that action as the manner in which the dwarf is trying to get the rat out of its hole.

- (59) [20170331c_n01s140_60]
 mi-sak pis-i va-leh rat asu va-mak
 3SG.NFUT-do try-3OBJ 3SG.IFUT-take out rat 3SG.IFUT-like.this
 '[He pulled and pulled and pulled...] he tried to take out the rat like this.'

Finally, *mak* can also have a recognitional use. In (60), the speaker uses *mak* to refer to the way in which young people often just stay home, assuming the audience's shared knowledge of the common social problem of young people hanging around without enough to do.

- (60) [20170406a_n01m164_61]
 yangfala xil, samtaem ra-di mak
 youth PL sometimes 1PL.INCL-NFUT.stay 3SG.NFUT.like.this
 'Young people, sometimes we stay like this.'

Because *mak* has a bilabial onset, it usually does not take the non-future third-person singular subject-indexing prefix *mi*, instead taking a zero allomorph. This means that it most often appears in its unprefixed form, as in all the examples here except (59). This bridging context seems to have allowed reanalysis as an invariant adverb. There are five unambiguous tokens in the corpus in tense and person contexts where prefixes are expected, but do not occur. However, in the non-future with a third-person singular subject it is usually not possible to disambiguate the verbal and adverbial lexemes.

5.3 Second-person proximal or contrastive *maxai*

While there is no evidence in the corpus or observation of any verbal forms based on distal *xor*, there is also one example of a verbal manner demonstrative *maxai* in the corpus, which appears to be the second-person proximal *xai* form. Example (61) is reported speech from the speaker's sister-in-law, instructing her (the addressee of the reported speech) not to roll simboro in the manner she was doing it, but instead in the manner the sister-in-law demonstrates. *Maxai* is the subsequent verb in the serial verb construction, modifying the initial verb phrase 'roll simboro'. While both verbs take negative polarity immediate future prefixes, the negative clitic *ti* only appears after the initial verb phrase, a strong diagnostic of serialisation, as *naa-maxai* would be ungrammatical as independent clause without the negative clitic.

- (61) [20141212g_p01s046_11]
- a-naa-pis siboro =ti naa-maxai e
2SG.IFUT-NEG-roll simboro =NEG 3SG.IFUT.NEG-like.that.PROX2 but
u-pis siboro i-xoni ngan =ak
2SG.DFUT-roll dumpling 3SG.DFUT-like one =PROX1
'You don't roll simboro like that, you roll it like this one.'

5.4 Verbal hesitation form *mue*

The verbal form *mue* is probably based on contrastive *-e*. Rather than functioning as a verbal manner demonstrative, this is a hesitation form used when a speaker is searching for a verbal lexeme, which is usually supplied by the speaker immediately afterwards. Alternatively a speaker might break off after *mue*, and use a different sentence structure altogether in the following sentence. It is very frequent, occurring roughly once every 210 words in the corpus (227 tokens).

Mue is fully inflected for the intended subject person-number and relative tense features, and can be modified by auxiliary verbs (64).

- (62) [20141208a_n01m045_25]
- mue, muis vatang
3SG.NFUT.HES 3SG.NFUT.cry a.lot
'He um, he cried a lot.'
- (63) [20150223a_n01m096_59]
- ral-mue, rali-pol
1DU.INCL.DFUT-HES 1DU.INCL.DFUT-work
'We'll um, we'll work.'
- (64) [20150303d_p01e016_03]
- di ma-mue, di ma-ga-ni
CONT 1PL.EXCL.NFUT-HES CONT 1PL.EXCL.NFUT-NFUT.eat-3OBJ
'We um, we eat it.'

Mue can stand in for both intransitive (62)–(63) and transitive (64) clauses. When a transitive clause is intended, it can also be followed by the nominal place-holder *neta* 'thing', see (51). This can be used whether the intended object will be indexed by the object pro-indexing suffixes (65), or an independent noun phrase (66).

- (65) [20141027a_n01m001_18–19]
 lu-mue neta, lu-pangas-i
 3DU.NFUT-HES thing 3DU.NFUT-burn-3OBJ
 ‘They um what, they burnt it.’

- (66) [20170331c_n01s140_15]
 asu mue neta, mi-gur vatupang
 rat 3SG.NFUT.HES thing 3SG.NFUT-NFUT.take fire.stick
 ‘The rat um what, took the fire stick.’

Mue has a bilabial initial and frequently appears as a bare stem in the non-future with a third-person singular subject, as in (62) and (66). Like *mak*, there are signs that this has been a bridging context for reanalysis as a non-verbal form. There are eight tokens in the corpus where a verbal reading is unlikely, and *mue* instead seems to be acting as a more general hesitation marker, with the syntactic distribution of an interjection.

Despite the formal similarity, and the evidence from Paamese of a deictic verbal root with an initial /mu/, it is difficult to assert a definite relationship between this form and the contrastive demonstrative suffix *-e*, because it does not function as a typical verbal manner demonstrative like *mak*. It is therefore important to ascertain a plausible path of reanalysis from the demonstrative uses of *-e* to the function of a hesitation marker. Himmelmann (1996: 234–235) observes that demonstratives are often used as fillers, and suggests that hesitation phenomena are closely associated with the recognitional uses of demonstratives. While *ak* is the primary strategy for recognitional use in Vatlongos, the complementary roles of *ak* and *-e* forms in discourse repair suggest a path for the verbal *-e* form to function as a hesitation marker. In examples like (51), *-e* modifies a placeholder for a lexical item that is currently inaccessible to the speaker, a close parallel to the role of verbal *mue*.

5.5 Summary

Table 10 summarises the functions of the verbal demonstrative forms. Here the imbalance in functional load between the first-person proximal series and the other forms is even more striking than in the adnominal, adverbial and pronominal forms discussed in §4. There is a single example of a second-person proximal verbal form, and the contrastive series is restricted to a hesitation form.

Table 10: Summary of functions of verbal demonstratives

Functions	<i>mak</i>	<i>maxai</i>	<i>mue</i>
referring to physical demonstrations by speaker	✓		
referring to manner of situational context	✓		
referring to manner in the preceding discourse	✓		
modifying manner of another verb	✓		
recognitional use	✓		
referring to manner of actions of addressee		✓	
hesitation			✓

6 Conclusion

There are four series of Vatlongos demonstratives: in addition to first-person proximal, second-person proximal and distal forms, there is a contrastive series that has not been previously described. These can be thought of as a person-based spatial paradigm, distinguishing first-person proximal from second-person proximal and distal series, and a contrastive paradigm where the first-person proximal series is distinguished from the contrastive series. The participation of the first-person proximal series in both paradigmatic distinctions is one reason for its greater frequency and wider discourse functions than the other demonstrative forms.

In discourse functions Vatlongos demonstratives are organised around a two-way distinction between forms based on the clitic *ak*, and contrastive forms based on the suffix *-e*. This chapter has described how asymmetries in the morphosyntactic status and distribution of these forms are reflected in their functional extensions. Whereas the first-person proximal clitic *ak* can freely modify both noun phrases and clauses, contrastive *-e* can only modify singular and plural noun phrases, usually in the third person, and its distribution is dependent on the host forms it attaches to, usually pronouns and quantifiers.

This asymmetry is mirrored in the more general and frequent contexts of use for the first-person proximal forms. First-person proximal *ak* forms are the default for anaphora and recognitional demonstrative uses, while *-e* forms are used for explicit contrast and negative affect. The negative affect meaning component is unusual in that it is not based on metaphorical distance: the primary meaning of these forms is contrastive rather than distance-based, so opposition seems to be a more relevant dimension for this connotation.

Vatlongos demonstrative verbs show an even starker asymmetry between these series. First-person proximal *mak* is by far the most frequent, used as a general verbal manner demonstrative, especially in serial verb constructions. There is also one example of second-person proximal *maxai*. *Mue*, the form probably based on contrastive *-e*, is restricted to a verbal hesitation form, reflecting the role of *-e* forms in modifying nominal placeholders. The role of the *ak* and *-e* series in discourse repair in nominal and verbal contexts is summarised in Table 11.

Table 11: The roles of *ak* and *-e* in discourse repair

Syntactic environment	<i>ak</i> series	<i>-e</i> series
nominal	<i>ak</i> modifies successful repair after hesitation	<i>xie</i> modifies placeholder for failed repair
verbal	<i>mak</i> refers to physical demonstrations	<i>mue</i> is a placeholder hesitation form during search for verbal lexeme

In summary, Vatlongos discourse demonstratives show how distinctions in semantic extension, discourse functions, and markedness interact with the morphosyntactic status and distribution of demonstrative forms, and the wider syntactic structures available in a language.

Abbreviations

In addition to Leipzig glossing rules, the following abbreviations are used:

COME	auxiliary of prior motion towards deictic centre	IFUT	immediate future relative tense
CONT	continuous (habitual or progressive)	NFUT	non-future relative tense
CONTR	contrastive	PART	partitive
DFUT	distant future	PC	paucal
GO	auxiliary of prior motion away from deictic centre	PRI	prior relative tense
HES	hesitation form	PROX1	first-person proximal (near speaker)
		PROX2	second-person proximal (near addressee)

References

- Aikhenvald, Alexandra Y. 2006. Serial verbs constructions in a typological perspective. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 1–68. Oxford: Oxford University Press.
- Barbour, Julie. 2012. *A grammar of Neverver*. Berlin: De Gruyter Mouton.
- Brotchie, Amanda. 2009. *Tirax grammar and narrative: An Oceanic language spoken on Malakula, North Central Vanuatu*. Melbourne: University of Melbourne. (PhD dissertation).
- Crowley, Terry. 1982. *The Paamese language of Vanuatu*. Canberra: Pacific Linguistics.
- Crowley, Terry. 1987. Serial verbs in Paamese. *Studies in Language* 11(1). 35–84.
- Crowley, Terry. 2002. *Serial verbs in Oceanic: A descriptive typology*. Oxford: Oxford University Press.
- de Vries, Lourens. 2005. Towards a typology of tail-head linkage in Papuan languages. *Studies in Language* 29(2). 363–384.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Early, Robert J. 1994. *A grammar of Lewo, Vanuatu*. Canberra: Australian National University. (PhD dissertation).
- Guérin, Valérie (ed.). 2019. *Bridging constructions*. Berlin: Language Science Press.
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntactic status of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin: De Gruyter Mouton.
- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Hyslop, Catriona. 2001. *The Lolovoli dialect of the North-East Ambae language, Vanuatu*. Canberra: Pacific Linguistics.
- Jauncey, Dorothy. 2011. *Tamambo, the language of west Malo, Vanuatu*. Canberra: Pacific Linguistics.
- Lichtenberk, Frantisek. 1996. Patterns of anaphora in To'aba'ita narrative discourse. In Barbara A. Fox (ed.), *Studies in anaphora*, 379–412. Amsterdam: Benjamins.
- Parker, Gary J. 1970. *Southeast Ambrym dictionary*. Canberra: Pacific Linguistics.
- Ridge, Eleanor. 2018a. Language contexts: Vatlongos, Southeast Ambrym (Vanuatu). *Language Documentation and Description* 15. 87–122.

4 Morphosyntactic and functional asymmetries in Vatlongos

- Ridge, Eleanor. 2018b. Vatlongos deposit. In Laboratoire de langues et civilisations à tradition orale (ed.), *Pangloss collection*. DOI: 10.24397/PANGLOSS-0000023.
- Ridge, Eleanor. 2019. *Variation in Vatlongos verbal morphosyntax: Speaker communities in Southeast Ambrym and Mele Maat*. London: SOAS, University of London. (PhD dissertation).
- Schneider, Cindy. 2010. *A grammar of Abma, a language of Pentecost Island Vanuatu*. Canberra: Pacific Linguistics.
- Spencer, Andrew & Ana Luís. 2012. The canonical clitic. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 123–150. Oxford: Oxford University Press.
- Thieberger, Nicholas. 2004. *Topics in the grammar and documentation of South Efate, an Oceanic language of Central Vanuatu*. Melbourne: University of Melbourne. (PhD dissertation).
- von Prince, Kilu. 2015. *A grammar of Daakaka*. Berlin: De Gruyter Mouton.
- Zwickly, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59(3). 502–513.

Chapter 5

Tracking and recognitional use of Kalamang *opa*: Demonstrative of cognitive accessibility

Eline Visser

Lund University

The Papuan language Kalamang has an adnominal demonstrative *opa*, which marks referents that are cognitively accessible either because they are common ground or because they are previously mentioned in the discourse. It is mainly an endophoric demonstrative which has tracking and recognitional uses, but it can be used exophorically as well. It can be used for both maintaining and reintroducing referents. The demonstrative *opa* is co-lexified with a time adverb *opa* ‘just now; earlier’. A grammaticalisation path cannot be established, but similar co-lexifications are found in Bunaq (Timor-Alor-Pantar) and Indonesian (Austronesian).

1 Introduction

Kalamang, a Papuan language of the West Bomberai family spoken in the east of Indonesia (ISO 639-3 code: kgv), has an adnominal demonstrative *opa*. It is typically used to refer to or establish common ground, or to indicate that the referent has been mentioned before. In (1), the referent *canam* ‘man’ is mentioned at minute 1:23. The same referent re-occurs about a minute later, at 2:25, and is then marked with *opa* to indicate it is the same man as at 1:23.

- (1) a. [stim24_1:23]
ma canam-at kon-i koluk
3SG man-OBJ one-NUM.ACC meet
'She meets a man.'



- b. [stim24_2:25]
ma canam **opa** me koluk
3SG man OPA TOP meet
'She meets that man (that I've mentioned before)'

Building on the last twenty years' advancements in the study of the semantics and pragmatics of demonstratives, I carry out an analysis of the functions of Kalamang adnominal *opa*. §2 provides an overview of the study of demonstratives, including an outline of exophoric and endophoric uses of demonstratives, and discusses notions such as attention direction, shared knowledge and accessibility. §3 discusses the Kalamang demonstrative system, focusing on the proximal and distal roots *wa* and *me*. In §4, the formal (i.e. morphosyntactic) behaviour of *opa* is discussed, and it is shown that *opa* is part of the adnominal demonstrative paradigm. §5, the core of this chapter, goes into detail about the behaviour of *opa*. Adnominal demonstrative *opa* is co-lexified with a time adverb *opa* 'just (now); earlier', which occurs at the edge of the clause. This co-lexification of time adverb and demonstrative, or similar functions, is also found in Indonesian and the Timorese language Bunaq, as discussed in §6. Conclusions are drawn in §7.

All data is drawn from the Kalamang corpus, which is based on field work by the author. The corpus contains spontaneous conversations and narratives (abbreviated "conv" or "narr"), stimulus-based conversations such as director-matcher tasks and stimulus-based narratives such as the *Pear Story* (abbreviated "stim") and elicitation (abbreviated "elic"). Examples from recorded natural speech are coded according to the genre and number of the recording and the time stamp of the utterance (for example, [conv8_2:03]), whereas elicited examples are coded according to the convention of [elic_dem] (where the second part refers to the topic of elicitation). All sources can be found using these codes in the Kalamang collection (Visser 2020).

2 Background

Demonstratives are forms that refer to an entity (a referent or an event) by locating it in space, time or discourse. Traditionally, demonstratives are thought of as basically encoding spatial distinctions (Anderson & Keenan 1985; Lyons 1977; Diessel 1999; Dixon 2003; Fillmore 1997; Kemmerer 1999),¹ with a differ-

¹Most of these authors recognise there is a bias in descriptive grammars towards analysing demonstrative systems as being spatial, e.g. Anderson & Keenan (1985); Lyons (1977); Dixon (2003); Diessel (1999).

ence between person-oriented and distance-oriented systems (e.g. Lyons 1977). Later work, like Özyürek (1998) and Burenhult (2003), offers broader and more detailed analyses of the discourse functions of demonstratives.

Endophoric demonstratives commonly include demonstratives that are used for introducing, identifying and tracking referents in discourse. They make reference within the text and are used to create textual cohesion (Halliday & Hasan 1976: 33; 57–76). Among other uses, Himmelmann (1996) distinguishes between tracking and recognitional uses of demonstratives. A demonstrative is used for tracking if it is used for established referents to help the addressee keep track of what happens to whom. Demonstratives in recognitional use are employed for unestablished referents, but refer to specific shared knowledge between speaker and addressee. In other words, the recognitional use signals to the addressee that the referent is common ground (“mutual knowledge, mutual beliefs, and mutual assumptions”, Clark & Brennan 1991: 127). In this chapter, I use the terms “common ground” and “shared knowledge” interchangeably.

Exophoric demonstratives refer to entities in the speech situation (Halliday & Hasan 1976: 33). Twenty years of research on exophoric demonstratives has shown that there is disagreement on the precise nature of demonstrative semantics. While Diessel (2006) maintains that distance is the semantic core of demonstratives, others have suggested that attention status (Özyürek 1998) or accessibility (Burenhult 2003) are the primary semantic properties encoded by demonstratives in some languages. In Jahai (Austroasiatic), some demonstratives can be analysed in terms of cognitive accessibility. These demonstratives refer to referents which have the addressee’s current or previous attention or knowledge (Burenhult 2003: 366), but they are strictly exophoric in use, relating to physical accessibility in the speech situation. Accessible referents are also likely to be reachable, approachable, perceptible and close. Although such physical characteristics play an important role in accessibility, Levinson (2018: 31) notes that it “is also conceptual and refers to whether a referent is or is not in the common ground or attention of speaker and addressee”. Accessibility is also noted by Hanks (2005; 2011) to be commonly marked by demonstratives. He stresses the non-physicality of the notion of accessibility, linking it to shared knowledge (“mutual knowledge” in his words), memory, and prior discourse (Hanks 2005), thus removing the term accessibility from the sphere of exophoric demonstratives and relating it to endophoric demonstratives, like Levinson (2018).

3 Overview of Kalamang demonstratives

Kalamang word order is SOV, with nominative-accusative alignment. Verbs are largely unmarked for tense, aspect and modality. The noun phrase is left-headed. Table 1 lists the demonstrative roots and their syntactic uses.

Table 1: Kalamang demonstratives and their syntactic use

Form	Pronominal	Adnominal	Identificational
<i>wa</i> proximal	+	+	+
<i>me</i> distal	+	+	+
<i>owa</i> far distal		+	+
<i>yawe</i> elevational ‘down’	+		+
<i>osa</i> elevational ‘up’	+		+
<i>opa</i> demonstrative		+	

Kalamang has two commonly used demonstrative roots: *wa* for proximal reference, and *me* for distal reference. These are exophoric demonstratives, referring to entities in the speech situation. A third root, far distal *owa*, is much more limited in use. *Wa* and *me* can be used pronominally, adnominally and identificationally. Pronominal demonstratives occur instead of nouns, adnominal demonstratives co-occur with nouns, and identificational demonstratives occur in copular and non-verbal clauses (Diessel 1999: 4). Adnominal demonstratives occur in the right periphery of the noun phrase. The same form is used in all functions, as illustrated in (2) to (4). The difference between *wa* and *wat* and *me* and *met* is that of subject versus object form, respectively.

(2) Adnominal

- a. [stim6_11:45]

ma	enem	wat-a	tu
3SG	woman	PROX.OBJ-FOC	hit

 ‘He hits this woman.’
- b. [conv12_8:15]

an	pulor-ca	met	parua-ta
1SG	betel.vine-2SG.POSS	DIST.OBJ	pluck-NFIN

 ‘I plucked that betel vine of yours.’

(3) Pronominal

- a. [conv3_1:59]

ki wat napaki-kin ye ge
 2PL PROX.OBJ use-IRR or not
 ‘Are you going to use this or not?’
- b. [conv12_20:42]

ma met-a sanggara
 3SG DIST.OBJ-FOC search
 ‘She searches for that.’

(4) Identificational

- a. [stim15_2:03]

wa me yartep
 PROX TOP sardine
 ‘This is a sardine.’
- b. [conv11_4:25]

naharen-un-a me
 leftover-3POSS-FOC DIST
 ‘That is his leftover.’

At the present stage of my work on Kalamang, the demonstrative system is analysed as distance-based. At several occasions, speakers have expressed that *wa* is reserved for referents close to the speaker, and *me* for referents at some distance from the speaker. However, this is mainly valid in a contrastive situation, where two referents are at different distances from the speaker and need to be contrasted (*this* is my phone, not *that*). As an illustration, consider (5). The example quotes the speech of a monkey whose fur is slowly turning white. The demonstratives *wa* and *me* refer to the same referent: the monkey’s head. Nevertheless, the speaker first chooses proximal *wa* and then, after using *wangga* ‘from here’, chooses distal *me*. It seems that the speaker wants to create a contrast between the point ‘here’ (the neck, the point up to where the monkey has become white), and the monkey’s head ‘that’, where his black fur is still visible. There is no literal distance here, but by using *me* a contrast between the neck and the head is created.

- (5) [narr19_16:04]

wa tok kuskap wangga sara me tok kuskap
PROX still black from.here go.up DIST still black

'This is still black, going up from here, that is still black.'

Wa can also be used adnominally with the first person pronoun, to contrast between speaker and addressee or another referent. Example (6) shows the first person pronoun followed by proximal *wa* and topic marker *me*, where the use of the first person creates a distance between himself, who is married to a mermaid in the sea, and the other villagers, who are married to human beings on land.

- (6) [narr26_6:59]

an wa me laut-ko an se kion
1SG PROX TOP sea.MLY-LOC 1SG IAM married.male

'As for me, in the sea I am already married.'

The demonstratives *wa* and *me* have object forms *wat* and *met*, which include the object marker *-at*. Both adnominal demonstratives modifying an object NP and pronominal demonstratives in object position are inflected with the object case marker, as illustrated in (2a) and (2b) for adnominal *wa* and *me*, and (3a) and (3b) for pronominal *wa* and *me*. Modified nouns, such as *enem* 'woman' in (2a), remain uninflected.

The roots *wa* and *me* are used to derive all kinds of related forms. With the lative and locative case suffixes *-ka* and *-ko*, respectively, the meanings 'to/from here/there' and 'here/there' are derived. These forms are fossilised as *wangga/mengga* and *watko/metko*, respectively. Other related forms are the adverbs *wandi* 'like this' and *mindi* 'like that'.

As mentioned earlier, there is a third demonstrative root *owa* which is used for referents that are (relatively) far away. It is much more limited in use than *wa* and *me*, and is not used contrastively with *wa* and *me*. *Owa* is often inflected with locative *-ko* or lative *-ka*, surfacing as a predicative form *owatko* 'over there' or *owangga* 'to/from over there'. An example is given in (7), where the speaker has been at a distant market to sell goods.

- (7) [narr25_0:45]

an owatko eba esnem an-at kon sarie
1SG FAR.DIST.LOC then man 1SG-OBJ one chase

'I was over there but a man chased me.'

This limited use of *owa* is the reason to gloss the forms *wa*, *me* and *owa* as proximal, distal and far distal, rather than proximal, medial and distal. There is a clear opposition between *wa* and *me*, with *owa* only being used in exceptional cases.

Kalamang also has two elevationals. *Yawe* means ‘down’ and *osa* ‘up’. These can be used predicatively and adnominally. An adnominal example with *yawe* ‘down’ is given in (8). A predicative example with the locative form *osatko* is given in (9).

- (8) [narr22_5:59]
 som **yawe=ba** daruar
 person down=FOC pull.out
 ‘The person down there pulled it out.’

- (9) [conv16_3:16]
 in tok **osatko**
 1PL.EXCL still up.LOC
 ‘We were still up there.’

4 Morphosyntactic behaviour of *opa*

Used at the edge of a clause, *opa* is a time adverb meaning ‘just (now)’ or ‘earlier’. This is illustrated in (10).

- (10) [stim21_0:01]
 Yeni **opa** ka kareng-at kona
 Y. just 2SG frog-OBJ see
 ‘Yeni, did you see the frog just now?’

In this chapter I focus on the demonstrative *opa*, which is used adnominally. A typical example is (11), where the noun *balun* ‘his dog’ is followed by *opa*.

- (11) [stim20_2:42]
 bal-un **opa** kome sara~sara
 dog-3POSS OPA look up~DISTR
 ‘His dog [that I mentioned before] looked up.’

Unlike the demonstratives *wa* PROX and *me* DIST, demonstrative *opa* cannot be used pronominally, and cannot be inflected for case. Another difference between *opa* on the one hand and *wa* and *me* on the other is that the latter two

have longer forms (identical to the roots in use and meaning) with a prefix *yu-* and/or a suffix *-ne*, while *opa* cannot take these affixes. *Opa* thus stands apart morphosyntactically from *wa* and *me*, and there are some difficulties determining which slot *opa* occupies. Because there is no clear proof that *opa* occupies its own slot, and because of its demonstrative function, I consider *opa* to be part of the demonstrative paradigm. These arguments are illustrated in the following.

It is difficult to determine whether *opa* occupies the same slot as *wa* and *me*, or whether it occupies the slot before it. Demonstratives *wa* and *me* can be followed by a topic marker *me*, a homonym of the distal demonstrative. This is illustrated in (12) and (13).

- (12) [narr40_13:11]
in kelek-ko-i koyet leng wa me sontum saerak
1PL.EXCL mountain-LOC-VLNK finish village PROX TOP person NEG.EXT
'After we had gone to the mountains, there were no people in this village.'
- (13) [narr22_1:03]
yuol me me nauwanona-te
day DIST TOP pack-IMP
'That day you pack!'

Like the proximal and distal demonstrative, *opa* can also be combined with *me*, as in (14).

- (14) [narr6_0:23]
ror-at kuru mia terus kewe opa me pi he usar
wood-OBJ bring come then house OPA TOP/DIST 1PL already build
'[We] bring the wood and then we already build that house.'

The question is whether one should analyse *me* in (14) as the distal demonstrative or as the topic marker.² One way to find out whether there is a separate slot for *opa* before the demonstrative slot is to look at combinations of *opa* with other demonstrative forms. The combination of *opa* with the proximal form *wa* is not found; there are no occurrences in the corpus of adnominal use of *opa wa*. There is one example with the combination *opa wa*, which I analyse as biclausal (15): time adverb *opa* 'earlier' closes the first clause, and the second clause consists of a pronominal proximal demonstrative *wa* with the tag *to* 'right'. (The intonation of this utterance is inconclusive.)

²Translation to Indonesian does not help: *opa me* is usually translated as (*yang*) *tadi itu* 'the/that one earlier'. *Itu* is the distal demonstrative, but can also be used as topic marker in Indonesian.

- (15) [stim6_5:09]

[kon me ma koup-te **opa**] wa to
 one TOP 3SG hug-NFIN earlier PROX right
 'This one he was hugging earlier, here, right?'

The combination of *opa* with *yuwa*, a longer variant of *wa*, has a specific meaning 'earlier today', and is used at the beginning or end of a clause as a time adverb. *Opa wa* cannot mean 'earlier today', suggesting that *yuwa* in *opa yuwa* has lost its demonstrative meaning. There are no examples in the corpus where adnominal *opa* is combined with another demonstrative, except for locative forms. However, because locative constructions are predicative, locative demonstratives take the slot of the verb. Constructions combining *opa* and a locative, such as in (16), can thus be analysed as N + Dem + Pred.

- (16) [narr26_15:28]

supkalingN **opa**_{Dem} metko_{Pred}
 eel OPA DIST.LOC
 'That eel is there.'

I conclude that although it is not unequivocally proven that adnominal *opa* occupies the same slot as the demonstratives *wa* and *me*, I certainly do not have enough proof to place it in its own slot. With its demonstrative functions, which will be illustrated in the next section, I will therefore analyse *opa* as being in the slot for demonstratives. Table 2 summarises the demonstrative forms discussed in this section.

Table 2: Pronominal and adnominal demonstratives (selection)

Pronominal		Adnominal	
<i>wa</i>	PROX	<i>wa</i>	PROX
<i>me</i>	DIST	<i>me</i>	DIST
		<i>opa</i>	OPA

5 Functional behaviour of demonstrative *opa*

Demonstrative *opa* is a common and versatile form. After proximal *wa*, *opa* is the second most common demonstrative in the corpus. To illustrate the uses of

opa, I present a fragment from a four-minute narrative in (17). This story is a re-narration of a video that the speaker has watched.³ In (17), it is translated into English, with all instances of *opa* (in this case all combined with topic marker *me*)⁴ left untranslated in place. The four referents that are referred to with *opa*, namely the woman, the vase, the man and the fish, are bold to ease their tracing.

(17) [stim24_0:00–1:25]

- a. pas **opa** me
woman OPA TOP
'That woman,'
- b. karena pi komeret me
because 1PL.EXCL see TOP
'because we saw,'
- c. pas me me mungkin berupa jim atau
woman DIST TOP maybe.MLY a.kind.of.MLY ghost or.MLY
'that woman is maybe a kind of ghost or'
- d. e sontum purba
HES person ancient.MLY
'an ancient person.'
- e. terus ma **gusi-un** kon-i tanggo
then 3SG vase-3POSS one-NUM.ACC hold
'Then, she is holding one of her vases.'
- f. kewe-un-ka ma marmar-te
house-3POSS-LAT 3SG walk-NFIN
'From her house she walks ...'
- g. ma kewe-un-ka kaluar-te **gusi-un** kon-i tanggo
3SG house-3POSS-LAT exit-NFIN vase-3POSS one-NUM.ACC hold
'She comes out of her house holding one of her vases.'

³The video is computer-animated and without spoken language, displaying people dressed in Japanese fashion. It shows a woman who goes for a walk holding a vase. The vase is accidentally shot by a man with a bow, who later comes to her house to apologise. The video is available at https://www.youtube.com/watch?v=iPrn_raZxEs.

⁴About a quarter of the occurrences of *opa* in the corpus are without topic marker *me*. It is not clear to me what the difference between *opa* and *opa me* is, but note that both *opa* and *opa me* occur with all kinds of referents: names, pronouns, possessed nouns, and nouns. I am aware that the analysis of topic marker may not be the most appropriate in all instances; see, for example, (22), where two referents are marked with *opa me* in the same clause.

- h. **gusi-un** **opa** me iriskapten ma marmar terus
 vase-3POSS OPA TOP white 3SG walk further.MLY
 'That vase of hers is white. She walks further.'
- i. **pulpul** **pararuoren** bunga~bunga
 butterfly fly flower.MLY~DISTR
 'A butterfly flies, flowers ...'
- j. **ma uriap-un-ka** marmar marmar marmar
 3SG street-3POSS-LAT walk walk walk
 'She walks and walks along her street'
- k. **sampai** **mindai** me bo-te raor-ko
 until.MLY like.that TOP go-NFIN middle-LOC
 'until she comes to the middle'
- l. **ma canam-at** kon-i koluk
 3SG man-OBJ one-NUM.ACC meet
 'She meets a man.'
- m. **canam** me pusir-un-bon
 man DIST bow-3POSS-COM
 'That man has a bow.'
- n. **canam** jenis sountum me ba weinun jim atau sountum
 man kind.of.MLY person TOP but also ghost or.MLY person
 jaman purba
 era ancient
 'The man is a person but also a kind of ghost or ancient person.'
- o. a **mindi** ma kome-ta **sor** per ne-ko tiri~tiri
 HES like.that 3SG look-VLNK fish water in-LOC swim~DISTR
 par-un-bon-ten
 fin-3POSS-COM-REL
 'Like that she watches the fish in the water that are swimming with
 their fins.'
- p. **sor** hias
 fish decorated.MLY
 'Multi-coloured fish.'
- q. **komet** me **sor** **opa** tiri ran mian-ten
 look TOP fish OPA swim go come-REL
 'Looking ... those fish are swimming back and forth.'

- r. mungkin ma met komet~komet **gusi-un** opa me
maybe.MLY 3SG DIST.OBJ look~DISTR vase-3POSS OPA TOP
tan-un-ka tur
hand-3POSS-LAT fall
'Maybe [while] she is looking at that that vase of hers falls from her hands.'
- s. **gusi-un** opa me tan-un-ka tur [...]
vase-3POSS OPA TOP hand-3POSS-LAT fall
'Her vase falls from her hands.'
- t. ma he marmar yecieni me raor-ko ma **canam** opa me koluk
3SG IAM walk return TOP middle-LOC 3SG man OPA TOP meet
'She walks back to the middle, she meets that man.'

The first use of *opa* in this fragment, with ‘woman’ in (17a), indicates to the addressee (me, the linguist) that she “knows” this woman, because she has also seen the video where the woman appears.⁵ That is, ‘woman’ is marked as being shared knowledge. This is an example of the recognitional use of *opa*.

The uses of *opa* with *gusi* ‘vase’ in (17h) and (17s), with *sor* ‘fish’ in (17q) with *canam* ‘man’ in (17t) are tracking uses, helping the addressee “in keeping track of what is happening to whom” (Himmelmann 1996: 226). With ‘vase’ in (17h) and ‘fish’ in (17q), *opa* is used to immediately maintain the referent. ‘Man’ in (17t), on the other hand, is referred to with *opa* only after having been absent from the narrative for a while, i.e. to reintroduce a referent. This also counts for a later occurrence of ‘vase’ in (17s).

A tracking example from another text is given in (18). The referent, *semen* ‘concrete’, is mentioned at minute 2:16, and its repeated use is marked with *opa* at minute 5:27.

- (18) a. [narr7_2:16]
mu he semen-at cetak
3PL already concrete-OBJ mould
'They already mould the concrete.'
- b. [narr7_5:27]
mu he semen **opa** koyal-te di=ra-n
3PL already concrete OPA mix-NFIN TER-go.up-N
'They already mixed the concrete and put it up.'

⁵The same speaker starts his narration of the *Frog Story* (Mayer 1969) the same way ('child *opa* is sitting in his house'), but there he is narrating *while* he sees the pictures of the story.

In (19), the referent Desi is in the outside world, in fact just outside the window of where the speaker and addressee are sitting.

- (19) [conv11_6:36]

Desi **opa** me yal~yal-te yawe
 D. OPA TOP paddle~DISTR-NFIN down
 'Desi is paddling down there.'

In the entire village of the speaker there is only one person named Desi, who, moreover, is the speaker's daughter. As such she does not use *opa* to disambiguate which Desi she's talking about. Even though Desi has come up several times in the same conversation from which (19) is taken, the speaker does not seem to employ a tracking use to indicate "that Desi we've been talking about", since there is only one Desi. Rather, she is foregrounding a referent to which both speaker and hearer have a relation. Here, the boundary between endophoric and exophoric reference is a bit vague. On the one hand one would like to say that this is the recognitional use of *opa*, referring to shared knowledge. On the other hand, the use of *opa* while the referent is just outside the window is reminiscent of the exophoric use of demonstratives, indicating entities within the speech situation. Perhaps the definition of recognitional demonstratives, given as being a kind of endophoric demonstrative, should be widened to include exophoric reference as well, at least in Kalamang. See also (20), where the referent *pulor* 'betel leaf' is marked with *opa* in a tracking kind of use (having been referred to by the other speaker a few turns earlier), but at the same time it is an exophoric demonstrative as the betel leaves are somewhere in the room where speaker and addressee are located. The speaker even points to the vague direction where she thinks the betel leaves are located.

- (20) [conv12_8:51]

pulor **opa** ka tu-ta me tama-ba-kadok os-kadok
 betel.leaf OPA 2SG pound-TA TOP where-FOC-side beach-side
 'That betel leaf that you were pounding, on which side is it, on the
 beach-side?'

The reason the speaker does not opt for a basic spatial demonstrative (proximal *wa* or distal *me*) is that she wants to indicate that she is referring to the betel leaves that have been mentioned before. The speaker also refers to the common ground, by specifying that the addressee has been pounding the betel leaves she is looking for.

Example (21) illustrates nicely how common ground can be established, or perhaps rather reinforced, in discourse by using *opa*.

- (21) [narr14_3:14]
inier **opa** Hadi **opa** to
2DU OPA H. OPA right
'We two, with Hadi, right.'

At first, the speaker marks *inier* 'we two' with *opa*, assuming the addressee knows who she means. Then she starts to doubt this (in fact, there are a few repairs between the first *opa* and *Hadi*): does the addressee know both referents? One is the speaker herself, and the other one is her brother Hadi. She specifies this by saying *Hadi opa* 'that Hadi whom you know', indicating (again) that there is common ground, that the addressee knows this Hadi. Hadi has not been mentioned in the story from which (21) is taken. Again, this example can be analysed as a recognitional use of *opa*.

As a last characteristic of *opa*, note that it can be used with two referents in the same clause. Example (22) is taken from a narration of the *Pear Story* (Chafe 1975). Both 'goat', which was recently introduced in the narrative, and 'avocados', which has been an important referent in the story from the beginning, are referred to with *opa*.

- (22) [stim31_0:51]
lek **opa** me afokat **opa** me kona
goat OPA TOP avocado OPA TOP see
'That goat sees those avocados.'

This double use of *opa* shows that a speaker can choose to foreground two referents at the same time that the addressee can trace back to earlier mentioned referents. Again, this is an example of tracking use.

Summarising, the demonstrative *opa* has both tracking and recognitional uses (Himmelmann 1996). It can be used for both maintaining and reintroducing referents. Typically, *opa* is used as an endophoric demonstrative, but sometimes it can also be used exophorically. The overarching characteristic of these uses of *opa* is cognitive accessibility in the sense of Hanks (2005), covering both shared knowledge and previous discourse. In all examples in this section, by using *opa* the speaker signals that the referent is accessible within the discourse because it has been mentioned before (tracking use) or accessible from the addressee's knowledge because the addressee has engaged with this referent in real life outside the discourse (recognitional use).

This contrasts with the use of proximal *wa* and distal *me*, where the main factor in the choice of demonstrative is decided by space. There are no examples of these demonstratives in the corpus where they are used to highlight that the referent is (or is not) common ground, or that the referent has (or has not) been mentioned before.

6 Co-lexification of demonstrative and adverb

Kalamang has two lexemes *opa*, which occur in different syntactic positions. At clause edges, it is a time adverb that translates as ‘just now’ or ‘earlier’. Adnominally, it is a demonstrative. The analysis of demonstrative *opa* as presented here, namely as a marker of cognitive accessibility, helps to understand how these two lexemes are related and how co-lexification might have arisen. What has happened recently is accessible within the minds of those who have witnessed it. It is easy to imagine a grammaticalisation path where the time adverb has come to be applied in a more abstract way as an adnominal demonstrative. In its tracking sense, pointing back to earlier mentioned referents, the use of *opa* is still quite literal. In its recognitional use, a generalisation seems to have taken place from a use of *opa* with referents that are accessible because they have been mentioned earlier to a use of *opa* with referents that are accessible because they are part of the common ground. However, there is no diachronic evidence to support this grammaticalisation path from time adverb to demonstrative.

Kalamang is not the only language displaying the co-lexification of these or very similar functions. Indonesian (Austronesian) uses *tadi* as an adjunct of time meaning ‘just recently’ or ‘a while ago’, but also adnominally “to refer back to someone or something recently mentioned or recently present” (Sneddon et al. 2012: 135). The Indonesian-English dictionary by Echols et al. (1989) gives both ‘a while ago’ and ‘aforementioned’ as translations of *tadi*. These meanings are illustrated in (23) and (24). Both glossing and translations are mine in these examples.

- (23) Indonesian (Austronesian; Echols et al. 1989: 540)
- | | | | | |
|-------------|-----|---------|----------|-------|
| tadi | ia | di.sini | sekarang | tidak |
| a.while.ago | 3SG | here | now | NEG |
- ‘He was here a while ago; now he is not.’

- (24) Indonesian (Austronesian; Indonesian translation of Paolo Coelho's *The Alchemist*; source: web)⁶

anak tadi mulai naik turun tangga~tangga istana
child aforementioned start go.up go.down stairs~PL palace
'The aforementioned child started to go up and down the stairs of the palace.'

Kalamang belongs to the three-language family West Bomberai. One other language in the family, Mbaham, has a demonstrative that seems to have a very similar function to demonstrative *opa*. In a very brief Indonesian grammar sketch, three demonstratives are listed (Flassy et al. 1984): *nggismo/kismo* seems to be proximal (translated as *ini* 'this'), *unggya* seems to be a distal (translated as *itu* 'that'), and a third form *ntoit* is translated as *tadi (itu)* 'earlier'. Unfortunately, no glossed examples are given. It is unclear whether *ntoit* is also an adverb of time. This could be an area of further study in that language.

Kalamang is possibly related to other languages of western New Guinea that have been hypothesised to belong to the Trans New Guinea macro-family. This includes Bunaq, a Timor-Alor-Pantar language of Timor (Usher & Schapper submitted). Bunaq has a word *mete* (now) that is functionally similar to both uses of Kalamang *opa*. *Mete* is analysed as a temporal/discourse locational (Schapper 2009: 285).⁷ When used independently, it refers to the present. Within the NP it is most commonly used to redirect attention to an established referent. It can also be used as a topic shifter, or, when used with a new referent, as a marker of a new stage in discourse. Schapper (2009) gives the following example from a narrative to illustrate the marking of an established referent. The referent, the monkey, has been absent for two utterances, and is reintroduced in (25) with *mete*. The NP is bracketed.

- (25) Bunaq (Timor-Alor-Pantar; Schapper 2009: 298)

hilaq [mete orel himo] g-ini g-ie
surprise NOW monkey CONTR.AN 3AN-CAUS 3-POSS

'What a surprise, it was this monkey just now who had made (the child) his.'

⁶<https://ndiethesmartass.wordpress.com/2010/01/22/rahasia-kebahagiaan/>

⁷Locutionals form a distinct word class from determiners (which include demonstratives) in Bunaq, because there are some syntactic and semantic differences between the two. Both classes, though, are used for locating and identifying referents. The distribution of Bunaq locational *mete* and Kalamang demonstrative *opa* does not seem to be very different.

Bunaq adnominal *mete* functions in a very similar way to Kalamang adnominal *opa*, while also being used as an indicator of time.⁸

The data from Kalamang, Indonesian and Bunaq suggests that the co-lexification of a referent tracker (a word of unknown word class in Indonesian, a locational in Bunaq, a demonstrative in Kalamang) and a time indicator is a phenomenon that might be worth investigating further, especially diachronically, to understand whether a grammaticalisation path can be discerned.

7 Conclusion

This chapter presented a careful analysis of the functions of Kalamang demonstrative *opa*. In contrast to proximal *wa* and distal *me*, which are spatial exophoric demonstratives that can be used pronominally and adnominally, *opa* is mainly an endophoric demonstrative that is only used adnominally. It is used for tracking referents that have been mentioned before, as well as for indicating that the referent is part of the common ground between speaker and addressee (recognitional use). Demonstrative *opa* thus signals to the addressee that the referent is cognitively accessible. Interestingly, demonstrative *opa* sometimes occurs outside strict endophoric use, referring to entities in the speech situation. In its tracking use, it can be used to maintain referents or to reintroduce them.

Demonstrative *opa* is homonymous with a time adverbial *opa* ‘just now; earlier’. There is no proof for a grammaticalisation path from one to the other, but it should be noted that Indonesian and Bunaq (Timor-Alor-Pantar) have co-lexified words with similar functions to Kalamang *opa*, suggesting a possible area for further study.

Abbreviations

The chapter follows the Leipzig Glossing Rules. Additional abbreviations include:

AN	animate	MLY	Malay loan
CONTR	contrastive	N	optional verb class
EXT	existential		marker of class ‘N’
FAR.DIST	far distal	NFIN	non-final verb
HES	hesitation	NOW	temporal/discourse
IAM	iamitive		locational
LAT	lative	NUM	number

⁸Note that Bunaq also has the words *meten* ‘before, past’ and *metensi* ‘just now’. It is unclear, however, whether these are related to *mete* (NOW).

TA	unanalysed suffix <i>-ta</i>	TER	terminative
TE	unanalysed suffix <i>-te</i>	VLNK	verb linker

Acknowledgements

Thanks to Niclas Burenhult, Åshild Næss and anonymous reviewers for their useful comments.

References

- Anderson, Stephen R. & Edward L. Keenan. 1985. Deixis. In Timothy Shopen (ed.), *Language typology and grammatical description*, vol. 3: *Grammatical categories and the lexicon*, 259–308. Cambridge: Cambridge University Press.
- Burenhult, Niclas. 2003. Attention, accessibility, and the addressee: The case of the Jahai demonstrative *ton*. *Pragmatics* 13(3). 363–379.
- Chafe, Wallace. 1975. *The pear film*. Berkeley, CA: University of California at Berkeley. <http://www.linguistics.ucsb.edu/faculty/chafe/pearfilm.htm>.
- Clark, Herbert H. & Susan E. Brennan. 1991. Grounding in communications. In Lauren B. Resnick, John M. Levine & Stephanie D. Teasley (eds.), *Perspectives on socially shared cognition*, 127–149. Washington, DC: American Psychological Association.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1). 61–112.
- Echols, John M., Hassan Shadily, John U. Wolff & James T. Collins. 1989. *Kamus Indonesia-Inggris: An Indonesian-English dictionary*. New York: Cornell University Press.
- Fillmore, Charles J. 1997. *Lectures on deixis*. Stanford, CA: CSLI.
- Flassy, Don A. L., Constantinopel Ruhukael & Frans Rumbraver. 1984. *Struktur Bahasa Bahaam*. Jayapura: Universitas Cenderawasih.
- Halliday, M. A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Routledge.
- Hanks, William F. 2005. Explorations in the deictic field. *Current Anthropology* 46(2). 191–220.
- Hanks, William F. 2011. Deixis and indexicality. In Wolfram Bublitz & Neal R. Norrick (eds.), *Foundations of pragmatics*, 315–346. Berlin: De Gruyter Mouton.

- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Kemmerer, David. 1999. “Near” and “far” in language and perception. *Cognition* 73(1). 35–63.
- Levinson, Stephen C. 2018. Introduction: Demonstratives: Patterns in diversity. In Stephen Levinson, Sarah Cutfield, Michael J. Dunn, N. J. Enfield & Sérgio Meira (eds.), *Demonstratives in cross-linguistic perspective*, 1–42. Cambridge: Cambridge University Press.
- Lyons, John. 1977. Deixis, space and time. *Semantics* 2. 636–724.
- Mayer, Mercer. 1969. *Frog, where are you?* New York, NY: Dial Press.
- Özyürek, Aslı. 1998. An analysis of the basic meaning of Turkish demonstratives in face-to-face conversational interaction. In Serge Santi, Isabelle Guaïtella, Christian Cavé & Gabrielle Konopczynski (eds.), *Oralité et gestualité : Communication multimodale, interaction*, 609–614. Paris: Harmattan.
- Schapper, Antoinette. 2009. *Bunaq: A Papuan language of central Timor*. Canberra: Australian National University. (PhD Thesis).
- Sneddon, James Neil, Alexander Adelaar, Dwi N. Djénar & Michael C. Ewing. 2012. *Indonesian: A comprehensive grammar*. London: Routledge.
- Usher, Timothy & Antoinette Schapper. Submitted. The Greater West Bomberai languages.
- Visser, Eline. 2020. *The Kalamang collection: An archive of linguistic and cultural material from Karas*. Lund: Humlab, Lund University Humanities Lab. <http://hdl.handle.net/10050/00-0000-0000-0003-C3E8-1@view>.

Chapter 6

Stsíkiistsi ki stsíkiistsi: The ubiquity of Blackfoot demonstratives in discourse

Heather Bliss

Simon Fraser University

Martina Wiltschko

Institució Catalana de Recerca i Estudis Avançats, Universitat Pompeu Fabra

Blackfoot demonstratives are ubiquitous and richly polysynthetic, yielding an inventory of 900 unique demonstrative forms. In oral stories, the distribution of demonstratives is particularly broad, and many appear with no clear referent or nominal complement, making no semantic contribution to the propositional content of the utterance. These “untranslatable” demonstratives are the focus of this chapter. Drawing on data from oral stories, we catalogue the properties of untranslatable demonstratives as a means to identify their discourse functions, and we demonstrate that different morphological and prosodic properties encode different discourse functions such as epistemic stance, noteworthiness, and emotivity.

1 Introduction

Demonstratives in Blackfoot (ISO 639-3: bla), a Plains Algonquian language spoken in Alberta and Montana, are ubiquitous, fulfilling a wide variety of syntactic functions as determiners, pronouns, predicates, and temporal expressions.¹ Demonstratives themselves are polysynthetic in Blackfoot, comprised of up to six morphemes, yielding 900 unique forms. In oral stories, the distribution of demonstratives is particularly broad; in addition to their aforementioned uses, they

¹The title of our paper, *Stsíkiistsi ki stsíkiistsi*, reflects the ubiquity of deictic elements in Blackfoot more generally. Translating loosely as ‘others and others’, it is comprised of conjoined and pluralised forms of the deictic root *stsíki*, which has no apparent morphological relationship to other demonstrative or pronominal paradigms.



can be used without any clear referent or nominal complement and without any truth-conditional contribution to the content of the utterance. An example of this use of the demonstratives is given in (1).² The demonstrative *mii* here does not refer to any discourse referent nor does it have any of the other functions typically associated with demonstratives, as we will show. It seems untranslatable into English.

- (1) (Beatrice Bullshields: *Itohkanao'tsisiyo'p*)
Mii ki oostówawáyi stámsaomatapooyaa.
am-yi ki o-isto-wa-ayi stam-sa-omatap-oo-yi-aawa
DEM-OBV CONJ 3-PRN-PROX-PRED just-out-start-go.AI-PL-3PL.PRN
'They started out.'

We refer to these demonstratives as "untranslatable" and we propose that these demonstratives function as discourse markers. Drawing on data from a corpus of nearly a hundred oral stories, we document which demonstratives can be used in which discourse contexts and we identify the morphological and prosodic properties that encode different discourse functions such as the knowledge states of the interlocutors (*epistemicity*), their emotional states (*emotivity*), as well as how significant the contribution is to the conversation in context (*noteworthiness*). As such, the primary contribution of this chapter is empirical. To our knowledge, this is the first paper that documents the properties of untranslatable demonstratives in Blackfoot, and by mapping morphological and prosodic features onto discourse functions we are laying the foundation for further analysis of Blackfoot demonstratives as discourse markers.

The chapter proceeds as follows. In §2, we outline the sources of our data and our methods for cataloguing demonstratives. In §3, we describe Blackfoot demonstratives in terms of their morphological composition and syntactic distribution. In §4, we narrow our focus to the untranslatable demonstratives and examine the mappings between morphological and prosodic properties and discourse functions. In §5, we provide further evidence for our proposal by showing that particles and demonstratives form a natural class with other units of language that can function as discourse markers. In §6, we conclude.

²Data drawn from oral stories are listed with the storyteller's name (if available) and the title of the story. Uncited examples come from the first author's field notes. Examples are represented in the Blackfoot orthography developed by Frantz (1978).

2 Sources of data and methodology

The primary source of data in this chapter is the Blackfoot Oral Stories Database,³ an ongoing project developed by the first author in collaboration with members of the Siksika and Kainai communities of Southern Alberta, Canada (see Bliss et al. 2019). At the time of writing, the collection consisted of 95 oral stories (comprising approximately 350 minutes of audio recordings) told by 21 storytellers. Not all of the stories are transcribed and not all of them are publicly available on the website. Most of the stories in the collection were curated in group storytelling sessions, in which storytellers gather to share stories with each other, often around a particular theme. Other sources of data include Russell & Genee's (2014) collection of eight stories, the Glenbow Museum's online collection of eight stories,⁴ and the first author's field notes (2003–2019), which do not contain stories but isolated sentences or short monologues elicited under specific discourse conditions. Unlike the database, all the stories in the two other story collections are fully transcribed,⁵ and our understanding is that these stories were recorded in more formal settings (i.e. a single speaker alone in a sound booth). Data from field notes is not analysed in terms of discourse properties but is used in this chapter to provide background information on the grammar of demonstratives and particles.

There are two limitations to using this corpus for investigating the discourse properties of demonstratives. First, because it consists entirely of monologic texts, the corpus includes no conversation data, which presumably would be the most fertile environment for observing discourse functions. However, as noted, many of the stories in the database were curated in group settings which tend to be rather informal and conversational in nature, and these show a marked difference from those curated in more formal contexts. We discuss this in more detail in §4.5 below. Second, stories from the Russell & Genee (2014) collection and the Glenbow collection are available as transcripts rather than audio recordings, and as such it is impossible to know whether there are any omissions from these transcripts that are relevant to this study, namely untranslatable material such as demonstratives.

Limitations aside, the corpus contains about 5100 demonstratives, which were catalogued according to their morphological, prosodic, distributional, and dis-

³ Available at <http://stories.blackfoot.atlas-ling.ca>.

⁴ Available at <https://www.glenbow.org/blackfoot/>.

⁵ A small number of the texts in the database include full or partial morphological analysis, whereas none of those in the Russell & Genee collection do. Four of the eight texts in the Glenbow collection have been analysed and glossed by the first author of this chapter.

course properties.⁶ For untranscribed audio recordings, translation assistance was provided by Siksika Elder and fluent speaker Ikino'motstaan Noreen Breaker. For the purpose of analysing untranslatable demonstratives, all demonstrative tokens with an obvious referent or nominal complement, and/or a clear semantic contribution to the propositional content of the utterance were set aside. All demonstratives with a “verbalising” suffix suggestive of a predicative function were also set aside. The remaining 530 forms comprise the corpus of untranslatable demonstratives; these were analysed for correlations between morphological/prosodic properties and discourse functions. These correlations are discussed in §4, following a brief introduction to the composition and distribution of demonstratives more broadly.

3 Blackfoot demonstratives: An overview

3.1 Composition of demonstratives

Demonstratives in Blackfoot are morphologically complex, comprised of an obligatory demonstrative root plus up to five optional suffixes, as illustrated in Figure 1 (descriptions of the roots and inflections are given in §4 below; see also Bliss (2013), Frantz (2017), Schupbach (2013) for detailed descriptions of the other morphemes, which are not of central importance to this chapter). There are no combinatoric restrictions on the composition of demonstratives, meaning that there are 900 possible demonstrative forms.

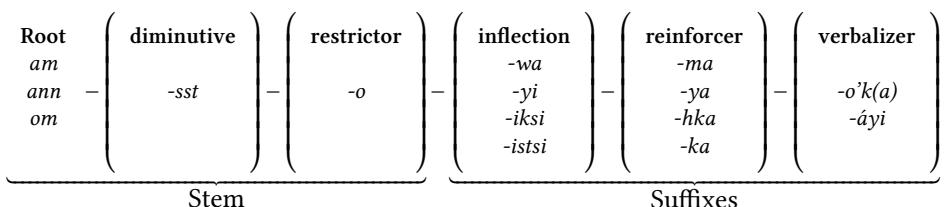


Figure 1: Blackfoot demonstrative template

In addition to these 900 forms, demonstratives can show variation based on prosodic properties, as described below.

Blackfoot is a pitch accent language (see, e.g., Van Der Mark 2003), and accent can be variably assigned to different syllables on the demonstratives. The

⁶In terms of prosodic properties, we catalogued vowel lengthening and reduplication. Pitch accent is also marked on the demonstratives but is yet to be catalogued and analysed. See §3.1 for discussion.

semantic and/or pragmatic contributions of pitch accent on the demonstratives are not yet well understood, but Schupbach (2013) observes correlations with syntactic function (i.e. different pitch accent patterns based on whether demonstratives function as determiners or pronouns) and speculates that there may also be specific pitch accent patterns for topicalised demonstratives. An analysis of the pitch accent patterns for the untranslatable demonstratives is pending, and as such pitch accent is not further discussed in this chapter.

Vowel lengthening and reduplication are also attested for demonstratives, the latter in particular with the untranslatable ones. These are discussed in §4.4. Finally, demonstratives are often combined with particles, particularly the conjunction particle *ki*. Particles are discussed in §5.

3.2 Distribution of demonstratives

Like demonstratives in many other languages, demonstratives in Blackfoot can function as either determiners or pronouns, but they also have a broader distribution, as will be observed throughout this section and in §4 below. Regarding their distribution as determiners, they are the only determiner-like elements in Blackfoot, and they have a canonical determiner-like distribution. They are required with arguments, as shown in (2), and are ungrammatical with predicate nominals (which we assume are pseudo-incorporated, see Bliss 2018), as shown in (3).

- (2) Kikatao'tsiksiiststoo'paatsiks *(omistsi) pisátssaisskiists?
 kit-kata'-otsiksiiststoo-'p-wa-atsiks om-istsi pisatssaisski-istsi
 2-INTERR-water.TI-1:INAN-PROX-3SG.PRN DEM-PL flower-PL
 'Did you water the flowers?'

- (3) Nitáíkskima (*oma) ponoká.
 nit-a-ikskimaa ponoka
 1-IMPF-hunt.AI elk
 'I am elk-hunting.'

In (2), the demonstrative *omistsi* cannot be omitted (as indicated by the asterisk outside the brackets), and in (3), the demonstrative *oma* cannot be included (as indicated by the asterisk inside the brackets). As pronouns, demonstratives can be interpreted as human or non-human, definite or indefinite, as shown in (4) through (6).

- (4) Nimáátowaanihpā ann.
nit-matt-waanii-hpa ann
1-NEG-say.AI-NONAFF DEM
'I am not saying that.'
- (5) Annoma ita'páíssiwa.
ann-o-ma it'ap-a-issi-wa
DEM-RESTR-STAT LOC-around-IMPF-be.AI-PROX
'She's around here.'
- (6) (Lena Russell: *Ksissta'pssiwa*)
Ísstsiiwoka ámohka awaasáííniwahka.
yisstsi-k-wa am-o-hka a-waasaii'ni-wa-hka
listen-IMP.PL-PROX DEM-RESTR-INVIS IMPF-cry.AI-PROX-REP
'Listen to someone crying'

In addition to functioning as determiners and pronouns, demonstratives can function as predicates. In this usage, they typically (but not necessarily) appear with a “verbalising” suffix, as shown in (7) and (8).

- (7) Óómahko'ka.
oom-wa-hk-o'ka
DEM-PROX-INVIS-VBZR
'It's way over there.'
- (8) (Glenbow: *Miohpokoaksi*)
Annimai akitaahsaopiiyaawa.
ann-yi-ma-ayi yaak-it-aahsa-opii-yi-aawa
DEM-OBV-STAT-VBZR FUT-LOC-always-stay.AI-PL-3PL.PRN
'And that is where they will always stay.'

Demonstratives can also be interpreted as temporal expressions in Blackfoot. Temporal demonstratives can have any of the syntactic functions described above: they can be determiners, pronouns, or predicates. Demonstratives are often (but not always) clause-initial when they receive a temporal interpretation, and they often (but not always) are marked with the “other time” suffix *-ka*, as in (9) and (10).

- (9) (Beatrice Bullshields: *Itáisapssiisskiitso'pa*)
 Amoka iiiskkóóhtsik nitsitsi'nakstssi'pi.
 am-o-ka isskoohtsik nit-it-i'naksstssi-hp-yi
 DEM-RESTR-OTH.TM long.ago 1-LOC-be.young.AI-CONJ.NOM-OBV
 'Long ago when I was young.'
- (10) (Glenbow: *Katoyissa*)
 Katoyissa anohk iitayo'kayihk omistsi Katoyissiksi.
 Katoyissa ann-o-hk ii-it-a-yo'kaa-ihk om-istsi
 K. DEM-RESTR-INVIS IC-LOC-IMPF-sleep.AI-REP DEM-PL
 Katoyissiksi
 Sweet.Pine.Hills
 'Katoyissa now sleeps at Sweet Pine Hills.'

In addition to their determiner, pronoun, predicate, and temporal uses, demonstratives are found with an even broader distribution in discourse contexts, particularly oral stories that are told in a conversation-like setting. These demonstratives do not have a clear referent or nominal complement, they do not contribute to the truth-conditional content of the utterance, and as such they cannot (easily) be translated. An example was given in (1), and a second example is given in (11).

- (11) (Beatrice Bullshields: *Itáisapssiisskiitso'pa*)
 Anna nódhkohpoksapopiiimook anna Tóotsinam.
 ann-wa noohk-ohpok-sapopii-m-ok ann-wa
 DEM-PROX please-ACCOMP-ride-TA-ACCOMP-2PL:3.IMP DEM-PROX
 Tóotsinam
 T.
 'Please ride with Tootsinam!'

The first instance of the demonstrative *anna* in (11) does not refer to an individual and it cannot be interpreted as either a predicate or temporal expression. Nor does it introduce a nominal complement; the second but not the first instance of *anna* functions as a determiner for the proper noun *Tóotsinam* (proper nouns in Blackfoot are often used with demonstrative determiners, see Bliss (2013) for examples and discussion). In short, the first instance of *anna* cannot be analysed as a determiner, pronoun, predicate, or temporal expression. Demonstratives of this variety are frequently found in oral stories, often clause-initially, and we refer to them as "untranslatable". We discuss their properties at length in §4.

4 “Untranslatable” demonstratives

4.1 Overview

In this section, we focus exclusively on the untranslatable demonstratives. As observed in §3.1, demonstratives vary along numerous morphological and prosodic dimensions. Here we look at which demonstratives are attested as untranslatable, and what the variation in untranslatable demonstratives can tell us about their discourse properties. We observe that a wide range of untranslatable demonstrative forms are attested, and that their morphological and prosodic variation correlates with properties typically associated with discourse markers, such as epistemicity, noteworthiness, and emotivity.

4.2 Demonstrative roots

When they function as determiners or pronouns, demonstratives are built from one of three roots, which are categorised according to a person-based proximity system, shown in Table 1.

Table 1: Demonstrative roots

Form	Meaning
<i>am-</i>	Proximity to speaker
<i>ann-</i>	Proximity to addressee
<i>om-</i>	Proximity to neither speaker nor addressee

All three of these roots are attested with untranslatable demonstratives⁷ but they appear in different discourse contexts. All of the *am-* and *ann-* forms appear in complementary distribution, with *am-* being used when the content of the proposition is something expected, predictable, or familiar to the addressee (12), and *ann-* being used when the content of the proposition is something unexpected, surprising, or new to the addressee (13).

⁷The initial vowel of untranslatable and other demonstratives is frequently dropped, obscuring the distinction between the *am-* and *om-* forms. In our corpus of untranslatable demonstratives, there are only three *om-* forms (with the initial vowel retained) and no *am-* ones. We assume that all of the *m-* forms are underlyingly *am-* based on their distribution.

- (12) (Clifford Crane Bear: *Who's in Charge?*)

Miiii náó'tapssapiwa otááksikamotaa'si.

am-yi na-o'tap-ssapi-wa ot-yaak-ikamotaa-hs-yi

DEM-OBV EVID-return-look.AI-PROX 3-FUT-give.birth.AI-CONJ-OBV

[A grandfather is trying to find someone who can transport his labouring wife to the hospital. After a few unsuccessful attempts, he returns home to check on his wife and then:] 'He returned to seeking help for her to give birth.'

- (13) (Mildred Three Suns: *Appapaainihkssin*)

Kiann nitáíto'too itohkanáópiyaa itáákopiyaa áikaahsiyaa.

ki-ann nit-a-it-o'too it-ohkana-opii-yaawa it-yaak-opii-yaawa

CONJ-DEM 1-IMPF-LOC-arrive.AI LOC-all-sit-3PL LOC-FUT-sit-3PL

a-ikaahtsi-yi-aawa

IMPF-gamble-3PL-3PL.PRN

[(Describing a dream:) The speaker hears music and is wandering around, trying to locate its source. Finally:] 'I arrived there and these people were all sitting, they would just sit there gambling.'

In (12), the demonstrative *miiii* is used in a sentence that reports a repeated event in the context of the narrative, namely the grandfather returning home yet again. As a repeated event, it is something that the audience (/addressee) could predict in the context of the narrative. Other instances of untranslatable demonstratives formed from the *am-* root have a similar distribution in narratives; they are used when the information is familiar or expected, often because it refers to repeated events in the story. Conversely, in (13) the demonstrative *kiann* is used at a point in the narrative when the speaker is recounting her discovery of something new and unexpected; it is surprising. In other narratives, untranslatable demonstratives formed from the *ann-* root are used with a similar distribution.

While the fact that the untranslatable demonstratives can function as markers of epistemicity status is not really surprising, the particular way in which they do this in Blackfoot is. The marking of epistemicity can be understood if we observe the fact that demonstratives are often used to mark the discourse status of referents. For example, in English (as in many other languages) demonstratives can be used to mark spatial deixis, indicating whether the referent is close or far from the deictic centre (*this* vs. *that*). However, spatial deixis is not the only dimension of contrast. Other common uses of demonstratives include anaphoric, discourse deictic, and recognitional uses (see Diessel (1999) for an extensive crosslinguistic

overview). Of particular interest to us is the discourse deictic use, which refers to the use of demonstratives as referring to propositions. This means that the demonstratives cannot only refer to individuals, but also to propositions. This is akin to the use of Blackfoot demonstratives under consideration here: they don't refer to individuals but instead they are used to express the epistemic state of the interlocutor relative to the proposition that hosts them.

Moreover, the recognitional use of demonstratives is also relevant: it refers to the use of demonstratives as indicating whether the referent is novel or familiar to the speaker, as in the English examples in (14). With the use of *this* in (14a), the speaker indicates that the referent is novel to the addressee; it is not part of the common ground. In contrast, in (14b) the use of *that* indicates that the speaker assumes that the addressee is familiar with the referent; it is part of the common ground.

- (14) a. A few years ago I saw *this* movie with Klaus Kinski.
b. Do you remember *that* movie we saw with Klaus Kinski?

Colasanti & Wiltschko (2019) demonstrate that this pattern is not restricted to English, but instead is found in many genetically and geographically unrelated languages, including Italian, German, Japanese, Medumba, Spanish, French, Southern Italo-Romance, and Cantonese.

Thus, for the Blackfoot demonstratives to indicate whether a given proposition is expected (/familiar) or unexpected (/novel) falls within this general use of demonstratives. However, there is one caveat: the Blackfoot pattern seems to be the exact opposite of what is observed in English and other languages. The root *am*, which can be used to indicate closeness to the speaker, would appear to be the equivalent to proximate demonstratives. But unlike proximate demonstratives in other languages, Blackfoot *am* indicates familiarity.

This unexpected pattern requires further research. In particular, it is necessary to carefully explore the use of the untranslatable demonstratives in conversations. Storytelling is a different kind of genre, where the interlocutors play different roles. For now, we speculate that a potential key to this conundrum might lie in the temporal use of Blackfoot demonstratives. In particular, recall from (9) and (10) that *am* is used to indicate temporal distance, whereas *ann* is used to indicate temporal proximity.

In sum, we have now seen that the untranslatable demonstratives can be used to mark the epistemic states of the interlocutors relative to the proposition. In this respect they carry a similar functional load to discourse markers in other languages (see Ariel 1998; Schiffrin 1987), including so-called “discourse particles” in

German (Thoma 2017; Zimmermann 2011) and Cantonese (Lam 2014), as well as discourse markers and enclitics in related Algonquian languages such as Ojibwe (Fairbanks 2016) and Meskwaki (Goddard 2015).

Following the hypothesis that the demonstrative roots encode epistemicity, we predict that the *om-* forms should also pattern together with respect to epistemic orientation, but this is not the case. The caveat, however, is that there are only three examples with *om-* and it is hard to draw reliable conclusions from such a small data set.

4.3 Demonstrative inflection

When they function as determiners or pronouns, demonstratives can take the same inflection as nouns. Blackfoot nominal inflection encodes number, animacy, and obviation distinctions, as shown in Table 2.

Table 2: Demonstrative inflection

		Animate	Inanimate
Singular	Proximate	- <i>wa</i>	- <i>yi</i>
Singular	Obviative	- <i>yi</i>	- <i>yi</i>
Plural		- <i>iksi</i>	- <i>istsi</i>

Whereas number inflection is relatively straightforward, animacy and obviation require brief explanations. Regarding animacy, all nouns in Blackfoot (and other Algonquian languages) are inherently categorised as animate or inanimate, but this grammatical designation does not always straightforwardly map onto ontological distinctions of animacy (i.e. not all animate nouns refer to what we tend to think of as animate individuals, at least under a Eurocentric worldview).⁸ Obviation is a reference-tracking device common to Algonquian languages; the more salient participant in a stretch of discourse is marked proximate and all subsidiary participants are marked obviative (see, e.g., Bliss 2017a; Dahlstrom 1991; Genee 2009; Goddard 1984, 1990; Junker 2004). Obviation is typically restricted to third persons in Algonquian languages, although first and second person independent pronouns are also marked with proximate and obviative inflection

⁸Some authors have suggested that Algonquian animacy distinctions may be motivated by ontological properties more congruent with an Indigenous worldview (such as “power potential”), suggesting that “animacy” is perhaps the wrong label for this category (see, e.g., Dahlstrom 1995, Goddard 2002).

in Blackfoot. As observed in Table 2, the proximate-obviative contrast is neutralised with inanimate and plural third persons; the syncretism between animate obviative and inanimate proximate/obviative can be considered accidental (see Bliss & Oxford 2016, 2017). Demonstrative determiners usually take the same inflection as the nouns they modify, with some exceptions (see Bliss 2017b on demonstrative-noun mismatches).

Whereas the full range of nominal inflection is attested when the demonstratives function as pronouns or determiners, only -wa (proximate singular) and -yi (obviative singular) are attested with the untranslatable demonstratives. Just as these inflections encode relative saliency when they appear on nouns and pronouns, they also appear to encode relative saliency when they appear on the untranslatable demonstratives. All of the proximate-marked demonstratives appear in sentences that coincide with salient or noteworthy events in the story (corresponding with, for example, plot twists or important points in the arc of the narrative), whereas all of the obviative-marked demonstratives appear elsewhere, without this same sense of noteworthiness. As such, the obviative-marked demonstratives are much more frequent, seemingly used as a default inflection. Representative examples of proximate and obviative forms are given in (15)–(16).

- (15) (Beatrice Bullshields: *Sa’aiwa*)

Kimaa ki ama nitómitaaminnaan mattsistáó’hkoonima omi
oyíiyiskaassini.

ki-am-wa ki am-wa nit-omitaa-m-innaan
CONJ-DEM-PROX CONJ DEM-PROX 1-dog-POSS-1PL

mattsista’-ohkooni-m-wa om-yi oyiiyis-kaa-ssin-yi
again-find.TI-3:INAN-PROX DEM-INAN nest-?-NOM-INAN

[A child is trying to protect some unhatched duck eggs, but the family dog repeatedly finds the nest and eats the eggs. This is the third and final time the dog does this before the child makes a drastic move to protect the one remaining egg:] ‘And then our dog found the nest again.’

- (16) (Clifford Crane Bear: *Isttsásápikimmaksi*)

Kiannika isáma’piiwa.

ki-ann-yi-ka isam-a’pii-wa
CONJ-DEM-OBV-OTH.TM long.time-be.II-PROX

[A grandfather is being questioned by the police and, as an aside, the storyteller notes that he had to wait for his grandfather to return:] ‘It was a long time.’

The demonstrative *kimaa* in (15) is formed from the root *am-* which, as we saw in the preceding section, functions to mark the utterance as containing familiar or expected information. In this example, as with many others, the information is expected because it is a repeated event in the narrative. However, although expected information, it is nevertheless noteworthy: this is a pinnacle point in the story which results in a major plot twist. The proximate marking *-wa* on the demonstrative foregrounds the event and signals to the audience that it is noteworthy or salient in the story. Conversely, the demonstrative *kiannika* in (16) is marked with the obviative suffix *-yi*, and the event denoted by the utterance is not one of particular significance in the story; it is simply an aside to the main event of the grandfather being held captive.

The function of noteworthiness is attested for discourse markers in other languages as well. For example, Jónsson (2017) notes that the Icelandic particles *en*, *vá*, and *æ* are required by some speakers with certain exclamatives to give them their noteworthy flavour.

4.4 Other demonstrative properties

The preceding sections looked at two sources of morphological variation in the demonstratives: root distinctions and inflectional distinctions. As observed in Figure 1, demonstratives exhibit other morphological variation as well. In addition to the inflectional suffixes, there are four other suffixes that can be added to the demonstrative roots. While some of these suffixes are attested with untranslatable demonstratives (see, e.g., (16) for the ‘other time’ suffix *-ka*), there are no detectable patterns of distribution in the corpus that would suggest these suffixes encode discourse functions.

However, vowel lengthening does appear to serve a discourse function with the untranslatable demonstratives. When the demonstratives are used deictically (i.e. as pronouns or determiners), vowel lengthening is iconic: the further away the referent is from the speaker, the longer the vowel is, as shown in the elicited examples in (17).

- (17) a. No lengthening

Na Carmelle	íinoyiiwa	omi	ksisskstakii.
ann-wa	Carmelle	ii-in-o-yii-wa	om-yi ksisskstaki-yi
DEM-PROX C.	IC-see-TA-DIR-PROX	DEM-OBV	beaver-OBV
‘Carmelle saw that beaver (nearby).’			

b. Slight lengthening

Na Carmelle íinoyiiwa oomi ksisskstakii.
ann-wa Carmelle ii-in-o-yii-wa om-yi ksisskstaki-yi
DEM-PROX C. IC-see-TA-DIR-PROX DEM-OBV beaver-OBV
'Carmelle saw that beaver (way over there).'

c. Pronounced lengthening

Na Carmelle íinoyiiwa ooomi ksisskstakii.
ann-wa Carmelle ii-in-o-yii-wa om-yi ksisskstaki-yi
DEM-PROX C. IC-see-TA-DIR-PROX DEM-OBV beaver-OBV
'Carmelle saw that beaver (waaaayyy over there).'

Vowel lengthening is also found with the untranslatable demonstratives and is sometimes quite pronounced, as in (18). In this example, the vowel of the demonstrative is extra-long with a duration of 0.877 seconds (compare this duration with that of the un-lengthened vowel of the conjunction particle *ki*, which is much shorter at 0.073 seconds).

- (18) (Clifford Crane Bear: *Isttsásápikimmaksi*)

Kimiiii nitáítopokská'sspiinaan.
ki-am-yi nit-a-itap-okska'si-hpinnaan
CONJ-DEM-OBV IMPF-toward-run.AI-1PL

[A young man is helping his grandfather with a task that the police later take notice of; he is confused about why they are doing what they are doing.] 'And we ran towards (something).'

What is the function of vowel lengthening when it appears on the untranslatable demonstratives? Our hypothesis is that vowel lengthening encodes emotivity, i.e. the speaker's emotive attitude or personal reaction – positive or negative – towards the content of the utterance. In all cases in which the vowel of an untranslatable demonstrative is lengthened, the context is one of heightened emotion within the narrative; consider (19) and (20).

- (19) (Beatrice Bullshields: *Itáísapssiisskiitso'pa*)

Kimiii nitóhtsitsinnohpíi nitsínnoo'takohpi.
ki-am-yi nit-oht-it-inn-ohpi'yí nit-inn-o'tak-ohpi'yí
CONJ-DEM-OBV 1-means-LOC-down-fall.AI 1-down-in.circle-fall.AI

[A child wants to go sledding with her sisters but they won't share their sleds. She finds a wash basin and gleefully slides down the snowy hill in the basin.] 'And with it, I went down, spinning in circles.'

- (20) (Beatrice Bullshields: *Itohkanao'tsisiyo'p*)
 Kimii nísto nisstónnatsotamasapaasai'ni.
 ki-am-yi n-iisto nit-sstonnat-sotam-sap-waasai'ni
 CONJ-DEM-OBV 1-PRN 1-extremely-genuinely-?-cry.AI
 [A child wants to accompany her grandparents to a ceremonial event, but they tell her she is too young and leave without her.] ‘I kept crying and crying after them.’

Both of these examples represent points in their respective stories in which the speaker has a strong emotional connection to the content of the utterance; in (19) this connection is positive and in (20) it is negative.

Emotivity is a discourse function encoded by discourse markers in other languages (see, e.g., Rett (2018) on emotive markers like English *alas*), and it has been observed to be achieved via prosodic features such as lengthening and pitch accent (see, e.g., Benus et al. (2007) on English *whatever*). Blackfoot exhibits a similar pattern: vowel lengthening on the untranslatable demonstratives correlates with instances of high emotion. As for (18), we suggest that the emotional connection is not with the speaker but with the speaker’s assumptions about the addressee; the storyteller is using vowel lengthening as a means to evoke suspense. All other examples in the corpus with vowel lengthening similarly coincide with strong emotions, either on behalf of the speaker or the addressee.

Whereas vowel lengthening is relatively common on demonstratives regardless of their syntactic function, demonstrative reduplication seems to be restricted to the untranslatable demonstratives only. Moreover, it is used frequently by one particular storyteller (Allan Stevens), less often by another (Mildred Three Suns), and not at all by the remaining 19 storytellers. Based on its limited distribution in the current corpus, it is not possible to ascertain the function of demonstrative reduplication, but we speculate that it may be used for emphatic purposes (21).

- (21) (Allan Stevens: *O'kaan*)
 Síksinnaki omi omi máátaanistsi'tooyi miksi matapíksi.
 siksinnaki-wa om-yi om-yi maat-waanistsi'to-yii-wa om-iksi
 police-PROX DEM-OBV DEM-OBV NEG-obey.TA-DIR-PROX DEM-PL
 matapi-iksi
 person-PL
 [At a Sundance ceremony, a police officer violated ceremonial protocol by crossing in front of some dancers, and later that day he fatally broke his leg. The moral of the story:] ‘The policeman didn’t listen to the people.’

In (21), the storyteller is concluding the story by emphasising the significance of the event and its consequences; this emphasis is encoded by the reduplication of the demonstrative.

4.5 Summary

In this section we have examined the composition and distribution of untranslatable demonstratives in Blackfoot, mapping their morphological and prosodic properties onto discourse functions. A summary of our hypotheses is presented in Table 3.

Table 3: Mapping demonstrative distinctions to pragmatic functions

Pragmatic function	Demonstrative distinction
Epistemicity	Root distinctions
Noteworthiness	Inflectional distinctions (proximate/obviative)
Emotivity	Vowel lengthening
Emphasis (?)	Reduplication

Table 3 shows that demonstratives in Blackfoot can encode a range of discourse functions that are typically associated with discourse markers in other languages. As such, we propose that Blackfoot demonstratives are multi-functional: they can fulfil canonical deictic functions as determiners or pronouns, but they can also fulfil discourse functions that extend beyond the truth-conditional content of the utterance. Given that untranslatable demonstratives are built using the same morphological and prosodic ingredients as all other demonstratives, we suggest that untranslatable demonstratives are not unique lexical items, homophonous with other demonstrative forms, but rather that all demonstratives have the capacity to function in various ways, including as discourse markers. In other words, demonstratives have the capacity to scope over individuals (as pronouns or determiners), times (as temporal expressions), situations (as predicates), or propositions (as discourse markers).

One outstanding question is whether the stories that serve as the data source here truly comprise a set of discourse contexts that are amenable to the use of discourse markers. As noted in §2, most of the stories in the database were curated in group storytelling sessions, which tend to be conversational in nature. To elaborate, in these group sessions, storytellers take turns sharing stories around a particular theme, and while the resulting recordings are indeed monological

rather than dialogical, the overarching context is much more like a dialogue or conversation with interlocutors exchanging information by sharing stories with each other. Importantly, it is only in texts from these conversation-like contexts that untranslatable demonstratives are found. More formal settings do not elicit untranslatable demonstratives, as evidenced by the fact that there are no examples of untranslatable demonstratives in the Russell & Genee (2014) collection or the Glenbow collection.

The properties of untranslatable demonstratives just reviewed suggest that they function in similar ways to discourse markers in other languages: they don't contribute to the propositional content of an utterance but instead serve to manage the ongoing conversation (e.g. Ariel 1998; Fairbanks 2016; Schiffrin 1987). In the following section we discuss further parallels between the untranslatable demonstratives and discourse markers.

5 Untranslatable demonstratives as discourse markers

Discourse markers are sometimes dedicated units of language (such as English *eh* and *huh*) and sometimes they are units of language that serve “double duty”: to express propositional content and to express discourse related functions (such as English *right*, *so* and *well*). However, what many discourse markers have in common is that they are often simplex forms which cannot be inflected. As such they are commonly labelled as “particles” (Abraham 1991; Aijmer 2002; Zimmermann 2011). While particles share certain formal properties (they are not complex, they don't inflect), they are generally not definable in positive terms and hence are not a natural class within and across languages. Nevertheless, it is perhaps not coincidental that discourse markers share these properties across different languages. Hence, in this section, we explore the properties of untranslatable demonstratives from the point of view of discourse markers.

By means of exploring the inventory of particles in Blackfoot, we show that particles that function as discourse markers are relatively rare in the language. At the same time, we show that they form a natural class with demonstratives. This lends further support to our claim that demonstratives in Blackfoot function as discourse markers.

The Blackfoot Dictionary (Frantz & Russell 1995)⁹ lists only 31 entries that are classified as UND (“underived particle”), excluding vocative pronouns and kin terms. Examples are provided in (22).

⁹There is a newer edition of the dictionary (Frantz & Russell 2017) but we have yet to conduct a search of this dictionary for UND forms.

(22) Blackfoot underived particles

- a. *saa* ‘no’
- b. *áa* ‘yes’
- c. *oki* ‘hello’
- d. *ki* ‘and’
- e. *tsimá* ‘where?’
- f. *tsa* ‘how?’ (often used in question formation)
- g. *yáoo ~ yáa* ‘oh no!’
- h. *ha’* ‘pfft! ha!’ (used to express scorn)

What these forms have in common is that they are underived particles and many of them function as discourse markers, i.e. they serve to regulate conversational interaction.

Some of the UND forms listed in the dictionary are concatenations of particles; see, e.g., (23).

(23) Concatenated particle

- ha’yaa
- ha’-yaa
- scorn-oh.no
- ‘uh-oh!’ (in anticipation of a reprimand)

Beyond these underived and concatenated particles, the majority of entries that are classified as UND in the dictionary involve demonstrative morphology in one of three ways. First, some inflected demonstratives are listed in the dictionary as UND forms, as in (24) and (25).

(24) Inflected demonstrative listed as UND in Frantz & Russell (1995)

- annóhk
- ann-o-hk
- DEM-RESTR-INVIS
- ‘right now’

(25) Inflected demonstrative listed as UND in Frantz & Russell (1995)

- ánniayi
- ann-yi-ayi
- DEM-INAN-VBZR
- ‘that’s enough’

Frantz & Russell's (1995) treatment of inflected demonstratives as underived particles suggests they are assuming them to be independent but homophonic lexemes. This differs from the analysis developed here under which inflected demonstratives are multifunctional and can take on a role as discourse markers. Our analysis is supported by the co-variation observed in §4 between the morphological and prosodic composition of demonstratives and their various discourse functions.

The preceding example shows one of the three ways in which entries classified as UND in the dictionary involve demonstrative morphology. The second way is one we have seen in earlier examples in this chapter: concatenations of a particle plus a demonstrative root (plus optional inflection). Additional examples are given in (26) and (27).

- (26) Particle + demonstrative listed as UND in Frantz & Russell (1995)

kiann
 ki-ann¹⁰
 CONJ-DEM
 'that's all' (often used to conclude a story)

- (27) Particle + demonstrative listed as UND in Frantz & Russell (1995)

há'annia
 ha'-ann-yi-ya
 PART-DEM-INAN-MVNG
 'Really! All right!'

Although particle-demonstrative concatenations, particularly with the conjunction particle *ki*, are rather ubiquitous in our corpus of untranslatable demonstratives, there is no clear pattern of distribution that suggests a particular discourse function can be assigned to particle+demonstrative combinations.

The third and final way in which demonstrative morphology is used to form what Frantz & Russell (1995) treat as underived particles involves a particle functioning as a base to which suffixes from the demonstrative paradigm can attach (with no demonstrative root); see (28) and (29).

- (28) Particle + demonstrative suffix listed as UND in Frantz & Russell (1995)

sááhkáyi
 saa¹¹-hk-ayi
 PART-INVIS-VBZR
 'just kidding'

¹⁰Compare: *ki* 'and'.

- (29) Particle + demonstrative suffix listed as UND in Frantz & Russell (1995)

kíka
ki¹²-ka
CONJ-OTH.TM
'wait!'

In sum, the majority of so-called particles listed in the Blackfoot dictionary are constructed using demonstrative morphology in some way. This suggests a formal connection between demonstratives and particles in Blackfoot, particularly because many of the demonstrative suffixes are restricted to only these classes of words (inflection is the one exception, as it appears on nouns as well). The fact that demonstratives and particles pattern together in terms of both their morphological composition and their ability to encode discourse functions, and the fact that the inventory of "true" particles is impoverished in Blackfoot, lends support to the proposal that demonstratives in Blackfoot function as discourse markers.

6 Summary and conclusions

Demonstratives in Blackfoot fulfil a wide variety of functions. They can scope over individuals (as determiners or pronouns), times (as temporal expressions), or situations (as predicates). This chapter focused on an additional function of the Blackfoot demonstratives which has not been previously documented: their ability to scope over propositions, as discourse markers. Demonstratives carrying out this role are "untranslatable"; they do not have a referent or nominal complement and they make no semantic contribution to the truth-conditional meaning of the utterance. Rather, their contribution is pragmatic: they encode the interlocutors' commitments, attitudes, and beliefs towards the propositional content of the utterance. As such, demonstratives in Blackfoot can function as discourse markers, and they can encode discourse properties typically associated with discourse particles in other languages.

Analogous to discourse markers in other languages, the primary contribution of Blackfoot untranslatable demonstratives is epistemicity, encoded by the demonstrative roots. The roots *am-* and *ann-* are in complementary distribution, with the former marking an utterance as familiar, expected, or repeated, and the latter marking it as unfamiliar, new, or surprising. Following Thoma (2017), we suggest that both of these encode an Addressee-oriented epistemic stance on the

¹¹Compare: *saa* 'no'.

¹²Compare: *ki* 'and'.

part of the speaker: they signal the Speaker’s expectations about the audience’s (/addressee’s) knowledge of the events denoted by the utterance.

In addition to the demonstrative roots encoding epistemicity, we have proposed that demonstrative inflection – and particularly the proximate/obviative contrast – is used to encode noteworthiness, with the proximate suffix signalling that the utterance is salient and deserving of the audience’s attention while the obviative suffix functions as the “elsewhere” case in opposition to the proximate.

Vowel lengthening signals emotivity (either positive or negative, of either the Speaker or the Speaker’s assumptions about the Addressee), and we speculate that reduplication signals emphasis.

Taken together, we observe that the wide range of morphological and prosodic variation in Blackfoot’s demonstrative paradigm allows for a wide range of discourse functions to be expressed by the demonstratives. Moreover, particles in Blackfoot are often concatenated with demonstratives or employ demonstrative morphology in their formation, suggesting that particles and demonstratives form a natural class, both with the capacity to function as discourse markers.

This connection between particles and demonstratives is not unique to Blackfoot. For example, some Dene demonstratives have grammaticalised as discourse particles (Thompson et al. 2011), and some German discourse particles are derived from Proto-Indo-European demonstratives (Thoma 2017). Synchronously, demonstratives are argued to take on other discourse functions such as negative appraisal in Hebrew and German (Sichel & Wiltschko 2018). While it is not unusual for demonstratives to take on the types of discourse functions discussed here, what sets Blackfoot demonstratives apart is their exceptionally rich morphology, and the ways in which the paradigmatic distinctions are used to encode a variety of different discourse functions.

Acknowledgments

Nitsiko’tahsi’taki and thank you to the Siksika and Kainai storytellers who have shared their words and wisdom. They are listed here in alphabetical order by English surname: Ahsikinaa Donna Axe, Mikiisistayii Shawn Axe, Apohyaki Trudy Axe, Dakota Back Fat, Iss-awa-waak-yaa Edith Breaker, Natóhkkitopi Fred Breaker, Naatoopii Lee Breaker, late Tootsinam Beatrice Bullshields, Clifford Crane Bear, Ikino’motstaan Noreen Breaker, Aahpiis Hanks Eagle Head, Pióhkomniaaki Rachel Ermineskin, late Annie Fox, Beverly Little Bear Hungry Wolf, Kah-tsim-kom-iki-kamosaahki Elsie Maguire, Roslyn Many Guns, Omahkaonopssi Dion Stevens, Sahsinaamskaa Allan Stevens, Kevin Stevens, Issookimiaaki Melinda

Stevens, Sohkapini Walter Stevens, Katai'tsinopaakii Mildred Three Suns, Evelyn Weaselhead, Ruth Yellowfly. We are also grateful to the following colleagues and advisors for their insights and support: Charm Breaker, Ikino'motstaan Noreen Breaker, Inge Genee, Marie-Odile Junker, Betsy Ritter, Leslie Saxon, Delasie Torkornoo, and audiences at the Workshop on the Discourse Functions of Demonstratives and the 50th Algonquian Conference. This work has been financially supported by the Jacobs Research Funds, the Phillips Fund, the Social Sciences & Humanities Research Council of Canada, and the University of Calgary.

Abbreviations

1, 2, 3	1 st , 2 nd , 3 rd person	NEG	negative
ACCOMP	accompaniment	NOM	nominal
AI	animate intransitive	NONAFF	non-affirmative
CONJ	conjunct	OTH.TM	other time
DEM	demonstrative	OBV	obviative
DIR	direct	PART	particle
EVID	evidential	PL	plural
FUT	future	POSS	possessive
IC	initial change	PRN	pronoun
II	inanimate intransitive	PROX	proximate
IMP	imperative	REP	reportative
IMPF	imperfective	RESTR	restricted
INAN	inanimate	SG	singular
INTERR	interrogative	STAT	stationary
INVIS	invisible	TA	transitive animate
LOC	locative	TI	transitive inanimate
MVNG	moving	VBZR	verbaliser

References

- Abraham, Werner (ed.). 1991. *Discourse particles: Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German*. Amsterdam: Benjamins.
- Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus*. Amsterdam: Benjamins.

6 The ubiquity of Blackfoot demonstratives in discourse

- Ariel, Mira. 1998. Discourse markers and form-function correlations. In Andreas H. Jucker & Yael Ziv (eds.), *Discourse markers: Descriptions and theory*, 223–259. Amsterdam: Benjamins.
- Benus, Stefan, Agustín Gravano & Julia Hirschberg. 2007. Prosody, emotions, and... ‘whatever’. In *8th Annual Conference of the International Speech Communication Association (INTERSPEECH 2007), Antwerp, Belgium, 27–31 August*, 2629–2632. https://www.isca-speech.org/archive/archive_papers/interspeech_2007/i07_2629.pdf.
- Bliss, Heather. 2013. *The Blackfoot configurationality conspiracy: Parallels and differences in clausal and nominal structures*. Vancouver: The University of British Columbia. (PhD dissertation).
- Bliss, Heather. 2017a. Dependencies in syntax and discourse: Obviation in Blackfoot and beyond. *Working Papers of the Linguistics Circle of the University of Victoria* 27(1). 1–26.
- Bliss, Heather. 2017b. Integrating demonstratives into Blackfoot’s nominal syntax. Invited colloquium at Memorial University of Newfoundland, April 2017.
- Bliss, Heather. 2018. Pseudo-incorporation in Blackfoot. *International Journal of American Linguistics* 84(4). 441–470.
- Bliss, Heather, Ilkino’motstaan Noreen Breaker & Elizabeth Ritter. 2019. *A’tsotsspommootsiiyo’p Niitsitapi’powahsin*: We help each other together with the Blackfoot language. Paper presented at the 6th International Conference on Language Documentation & Conservation (ICLDC): *Connecting Communities, Languages & Technology*, University of Hawai’i at Mānoa, February 28–March 3, 2019.
- Bliss, Heather & Will Oxford. 2016. A microparametric approach to syncretisms in nominal inflection. In Kyeong-min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson & Lisa Shorten (eds.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, 67–76. Somerville, MA: Cascadilla Proceedings Project.
- Bliss, Heather & Will Oxford. 2017. Patterns of syncretism in nominal paradigms: A pan-Algonquian perspective. In Monica Macaulay & Margaret Noodin (eds.), *Papers of the Forty-Sixth Algonquian Conference*, 1–18. East Lansing, MI: Michigan State University Press.
- Colasanti, Valentina & Martina Wiltschko. 2019. Spatial and discourse deixis and the speech act structure of nominals. Paper presented at the Canadian Linguistic Association Annual Meeting (CLA), Vancouver, June 1–3, 2019.
- Dahlstrom, Amy. 1991. *Plains Cree morphosyntax*. New York: Garland.

- Dahlstrom, Amy. 1995. Motivation versus predictability in Algonquian gender. In David H. Pentland (ed.), *Papers of the 26th Algonquian Conference*, 52–66. Winnipeg: University of Manitoba Press.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Fairbanks, Brendan. 2016. *Ojibwe discourse markers*. Lincoln, NE: University of Nebraska Press.
- Frantz, Donald G. 1978. Abstractness of phonology and Blackfoot orthography design. In William McCormack & Stephen A. Wurm (eds.), *Approaches to language: anthropological issues*, 307–325. The Hague: Mouton.
- Frantz, Donald G. 2017. *Blackfoot grammar*. 3rd ed. Toronto: University of Toronto Press.
- Frantz, Donald G. & Norma Jean Russell. 1995. *Blackfoot dictionary of stems, roots, and affixes*. 2nd ed. Toronto: University of Toronto Press.
- Frantz, Donald G. & Norma Jean Russell. 2017. *Blackfoot dictionary of stems, roots, and affixes*. 3rd ed. Toronto: University of Toronto Press.
- Genee, Inge. 2009. What's in a morpheme? Obviation morphology in Blackfoot. *Linguistics* 47(4). 913–944.
- Goddard, Ives. 1984. The obviative in Fox narrative discourse. In William Cowan (ed.), *Papers of the Fifteenth Algonquian Conference*, 273–286. Ottawa: Carleton University.
- Goddard, Ives. 1990. Aspects of the topic structure of Fox narratives: Proximate shifts and the use of overt and inflectional NPs. *International Journal of American Linguistics* 56(1). 317–340.
- Goddard, Ives. 2002. Grammatical gender in Algonquian. In H. C. Wolfart (ed.), *Papers of the 33rd Algonquian Conference*. Winnipeg: University of Manitoba Press.
- Goddard, Ives. 2015. The twenty-nine enclitics of Meskwaki. In Monica Macaulay & J. Randolph Valentine (eds.), *Papers of the Forty-Third Algonquian Conference*, 72–116. Albany, NY: SUNY Press.
- Jónsson, Jóhannes Gísli. 2017. Discourse particles and *hvað*-exclamatives. In Josef Bayer & Volker Struckmeier (eds.), *Discourse particles: Formal approaches to their syntax and semantics*, 100–114. Berlin: de Gruyter.
- Junker, Marie-Odile. 2004. Focus, obviation, and word order in East Cree. *Lingua* 114(3). 345–365.
- Lam, Zoe Wai-Man. 2014. A complex ForceP for speaker- and addressee-oriented discourse particles in Cantonese. *Studies in Chinese Linguistics* 35(2). 61–80.

6 The ubiquity of Blackfoot demonstratives in discourse

- Rett, Jessica. 2018. The semantics of emotive markers and other illocutionary content. Unpublished ms. Berkeley, CA: University of California at Berkeley. https://linguistics.ucla.edu/people/Rett/Rett_2019_emotive_markers.pdf.
- Russell, Lena & Inge Genee. 2014. *Ákaitśinikssiitsi: Blackfoot stories of old* (First Nations Language Readers Memoir 3). Regina: University of Regina Press.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Schupbach, Shannon Scott. 2013. *The Blackfoot demonstrative system*. Missoula, MT: University of Montana. (MA thesis).
- Sichel, Ivy & Martina Wiltschko. 2018. Demonstrative pronouns and the linguistic encoding of appraisal. In Will Bennett, Lindsay Hracs & Dennis Ryan Storoshenko (eds.), *Proceedings of the 35th West Coast Conference on Formal Linguistics, April 28–30, 2017, University of Calgary*, 365–373. Somerville, MA: Cascadilla Proceedings Project.
- Thoma, Sonja. 2017. *Discourse particles and the syntax of discourse evidence: Evidence from Miesbach Bavarian*. Vancouver: The University of British Columbia. (PhD dissertation).
- Thompson, Chad, Christopher Cox & Sally Rice. 2011. The grammaticalization of Dene demonstratives. Paper presented at the Dene Languages Conference, Whitehorse, Yukon, June 28, 2011.
- Van Der Mark, Sheena. 2003. *The phonetics of Blackfoot pitch accent*. Calgary: University of Calgary. (MA thesis).
- Zimmermann, Malte. 2011. Discourse particles. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft 33.2), 2012–2038. Berlin: De Gruyter Mouton.

Chapter 7

Pilagá determiners and demonstratives: Discourse use and grammaticalisation

Doris L. Payne

University of Oregon & SIL International

Alejandra Vidal

Universidad Formosa & CONICET

Guaykuruan languages of the South American Chaco have rich sets of classifiers and demonstratives, marking deictic, visibility, postural, and recognitional meanings. There is lack of consensus in the Guaykuruan literature about determiner and demonstrative elements, even across closely related dialects. This chapter explores them in Pilagá, including their structure, discourse profile, extension into the tense-evidentiality domain, and grammaticalisation as subordinators. Corpus data show that *mʔe* is best viewed as ‘distance neutral’, contrasting with *ho?* ‘proximal’ (which also has adverbial uses), *tfaʔa* ‘distal visible’, *maʃa* ‘unseen’, and *naqae* ‘recognitional’. *Mʔe* is dominantly endophoric and has grammaticalised as a relativiser. The ‘vertical’ classifying determiner *da?* has grammaticalised as a general subordinator.

1 Introduction

Pilagá (ISO 639-3: plg) is an endangered Guaykuruan language, spoken by around 5,000 people in Formosa, northeastern Argentina, in the South American Gran Chaco.¹ Guaykuruan languages have rich sets of determiners. Nearly all nouns require one, so they are ubiquitous in discourse. Pilagá determiners include forms that also function demonstratively, pronominally, or adverbially. They can have incipient nominal tense and evidential functions, and two have developed clausal

¹ Authors are listed in alphabetical order and the chapter is fully co-authored.



subordination functions. Some determiners are simple, involving only what we call classifiers (CLF),² highlighted in (1); others are demonstrative word-level constructions, highlighted in (2).³

- (1) (190Verbos2 165)⁴

s-anem **he-?**n nsedañanañat **da?** ya-qaya-di-pi
A1-give M-CLF:near pole CLF:VER POS1-brother-PAUC-COL
'I give the pole to my brothers.'

- (2) (Vidal 2001: 123)

ha-da=ča-lo yawo-?
F-CLF:VER=DEM1:DIST.VIS-PL woman-PAUC
'those women (standing)'

This chapter addresses the morphosyntax, meaning, and discourse uses of simple and complex Pilagá determiners and demonstratives.⁵ The study is based on a corpus of over 70 texts plus elicited data. §2 discusses definitions and terminology, and presents the three paradigms of key morphemes that figure in determiner and demonstrative constructions. Sections 3 through 6 focus on the morphosyntax and semantics of the constructions, supporting the claim that three distinct paradigms of key morphemes are involved. §7 discusses extensions into tense-evidentiality, and grammaticalisation of both the 'neutral' demonstrative root *m?e* as a relativiser and of the 'vertical' CLF *da?* as a more general subordinator. Throughout, issues of semantics and function are addressed, including how interaction among morphemes may affect interpretation. A conclusion is in §8.

²Classifiers are usually pro- or enclitics. We write them with the clitic boundary = as part of demonstrative constructions, but as separate orthographic words before a noun in accord with Pilagá orthographic practice. Some nouns with possessor prefixes lack determiners, though they can co-occur.

³Examples use a modified IPA representation with <y> for IPA /j/, <ñ> for /ɲ/, <č> for /tʃ/, for the bilabial fricative allophone of /w/; <λ> represents a palatal lateral sonorant. These are adaptations to the practical orthography.

⁴All data were collected by Alejandra Vidal with Pilagá native speakers in Formosa, between 1988 and the present. Data citations like "190Verbos2 165" refer to line "165" in file or text number "190" in our Pilagá FLEX database. The database contains narrative and expository texts, and some elicited material. Examples with no citation are elicited and not in the database.

⁵Previous studies of determiners and demonstratives in Guaykuruan discourse have focused on Toba, especially Carpio (2012) on Western Toba (which Vidal assesses as very close to Pilagá), González (2015) on an eastern variety of Toba, and Messineo & Cúneo (2019) on Toba generally.

2 Classifier and demonstrative roots

There is lack of consensus in the Guaykuruan literature about what are called “demonstratives”. The issues concern terminology for cognate forms and the inventory of relevant elements and complex structures, which may vary by dialect and language (Vidal 1997; 2001; Carpio 2012; González 2015; Messineo et al. 2016; Cúneo 2016). We thus first clarify key terms as used in this work.

- CLASSIFIER (CLF): Any of the six deictic or posture/shape clitics in Table 1.
- DEMONSTRATIVE ROOT (DEM1, DEM2): Any of the morphemes in Table 2 and Table 3, which have deictic, pointing-out, or joint-attention functions.
- DEMONSTRATIVE CONSTRUCTION (DEM): A word-level construction that contains a deictic or joint-attention establishing root **other than** just a classifier. All but one demonstrative construction contain a classifier; may function adnominally, pronominally, and in one case adverbially; and may be endophoric or exophoric to the discourse.⁶
- DETERMINER (DET): Any classifier or demonstrative construction **when functioning adnominally**. All determiners syntactically allow the noun (phrase) they accompany to function as a syntactic argument and/or as a referring expression in discourse. They may or may not be deictic.⁷

As the first two bullet points above suggest, we distinguish what we call classifiers (CLFS) from two sets of demonstrative roots (Table 1–3). Simple CLFS are the default determiner form in discourse (§3). Aside from a demonstrative construction with adverbial function (§4), all demonstrative constructions include a CLF. What we call DEM1 roots are preceded by a CLF (§5), while the DEM2 root is followed by a CLF (§6).

All the sets in Table 1–3 have some deictic semantics, and CLFS and DEM1 roots include visibility contrasts. The deictic overlaps might lead one to consider all three sets to be demonstrative morphemes. But there are functional reasons to distinguish classifiers from demonstrative roots. It would be unusual for a language to require every nominal to have a demonstrative, and this is one reason

⁶For Pilagá, we use “demonstrative root” to designate a root from Table 2 or Table 3, and “demonstrative” to designate a demonstrative construction.

⁷Diessel (1999: 57) uses “demonstrative determiner” for adnominal demonstratives not found in other syntactic contexts. In Pilagá, all demonstratives that function adnominally can also function pronominally. CLFS also (but rarely) function pronominally.

not to consider the default and ubiquitous classifier determiners to be demonstratives. The two sets of demonstrative roots are stronger orienting devices, pointing the hearer's attention to a participant, place, or time, usually (but not always) via deixis or visibility features.

Table 1 presents the Pilagá singular classifiers. Underlyingly they contain glottals, but they often surface with weak to no glottalisation. Classifiers with /a/ and /i/ often undergo vowel harmony alternation to /o/. For instance, *di?* has allomorphs *dyo* and *do?*.⁸ *So?* may undergo vowel harmony to *sa?*, and sometimes we find *eso?* We write the variations where they surface. The plural counterparts lengthen the vowel (Vidal 2001), though this is optional (especially when there is a plural affix on a noun).

Table 1: Pilagá singular classifier (clf) clitics

Deictic direction/Visibility		Posture/Shape	
<i>na?</i>	'near'; 'coming' to the reference point	<i>di?</i>	'horizontally extended' (line or plane)
<i>so?</i>	'far'; 'departing' from the reference point; 'past'	<i>da?</i>	'vertically extended'; 'abstract'
<i>ga?</i>	'unseen, absent'; 'unknown, generic, non-referential'; 'irrealis/future'	<i>ñi?</i>	'non-extended, bunched up, sitting'

As Table 1 shows, the clf paradigm has two semantic subsets (they are not contrastive in morphosyntactic distribution). Guaykuruan cognates of these morphemes have fascinated scholars due to the relatively unusual combination of their meanings, both basic and metaphorical (Klein 1979; Messineo et al. 2016). Relative to the physical world, the first semantic subset has deictic and/or visibility features, and in some contexts allows inference of motion semantics. The deictic meanings fit with Diessel's (1999) definition of demonstrative elements, which leads some researchers to refer to all Table 1 morphemes as "demonstratives" for related languages (Messineo et al. 2016).⁹ Carpio (2012) refers to the Western Toba cognates as "demonstrative roots" (she also identifies a distinct set

⁸Also note that *do(?)* is a dialect variant of *da?* 'vertical; abstract'.

⁹Messineo et al. (2016) do not mention cognates of the demonstrative roots we present in Table 2 and Table 3.

of morphemes – cognate with what we call “demonstrative roots” – to which the CLF cognates can attach). González (2015: 153) rejects calling the cognate Eastern Toba morphemes “classifiers” because, though they communicate a certain kind of nominal classification, a given nominal can occur with one or another according to the speaker’s perspective. (See also discussion in Messineo (2003: 145), who nevertheless uses the term “nominal classifier”.) However, in many classifier languages, classifier choice can be sensitive to varying speaker conceptions of the configuration of a concept. In Yagua, for example, ‘water’ can be conceptualised as long+horizontal or as round; ‘wood’ can be conceptualised as upright or as short+small (Payne 1986).

The second semantic subset most concretely indicates salient shape or postural configuration of a referent. For instance, *da?* in its concrete sense indicates vertically extended items like upright trees and people. It is also used for abstract nouns and has grammaticalised as a general subordinator (§7.3). The shape semantics lead Klein (1979), Vidal (1997; 2001), and Messineo & Cúneo (2019) to call all six “classifiers”. Our primary point here is not to argue that these six morphemes are (not) classifying or are (not) deictic in nature as the paradigm clearly has both types of semantic features. Rather, we wish to clarify that these comprise a distinct paradigm from what we call “demonstrative roots”, to which we now turn.

Pilagá demonstrative roots divide into two paradigms based on how they combine with classifiers: DEM1 roots follow CLFS, but the DEM2 root precedes CLFS. The DEM1 roots are deictic, indicating ‘proximal’, ‘distal’, and possibly ‘medial’ distinctions relative to a reference point, as well as visibility contrasts. To give an initial sense of their differing discourse profiles, Table 2 and Table 3 show the frequency of each demonstrative root in the corpus (whether as part of complex demonstrative constructions or not).

The DEM1 roots can function exophorically. This is most common for *ho?* ‘proximal’, which is dominantly exophoric and neutral for visibility, and for *ča?*a (variant *č?*a) ‘distal visible’.¹⁰ The exophoric uses optionally occur with pointing gestures. In contrastive elicitation contexts, *m?*e can indicate exophoric referents medially distant between those marked with *ho?* and *ča?*a. *Ma?*a refers to something unseen; the referent may be inferred or something about it is uncertain.

Demonstratives with *ho?* can also function endophorically to refer to discourse-

¹⁰For Western Toba, Carpio (2012: 47–49) identifies *-ha* ‘non-visible exophoric’ as a suffix on what we call CLFS. A possibly cognate Pilagá form surfaces in the frozen combination *ho? daha* ‘there, a place very far away’. We do not treat *-ha* further here but note its analogous position to *ho?*. *Ča?*a (variant *č?*a) comes from a motion verb and sometimes carries ‘itative’ and ‘ventive’ directionals that are characteristic of verbs, as in (38) and (39).

Table 2: Pilagá deictic and visibility demonstrative roots (DEM1)

DEM1 Root	Major Senses	Instances in Corpus
<i>ho?</i>	exophoric adverbial; ‘proximal’ (PROX); current discourse topic	363
<i>m?e</i>	exophoric ‘medial visible’; endophoric ‘neutral’ (NEUT)	241
<i>ča?a</i>	‘distal visible’ (DIST.VIS)	18
<i>ma?a</i>	‘unseen (NVIS)’; ‘inferential, uncertain’	6

anaphoric referents. Endophoric use is also possible for *ča?a*, but the referent is not considered close to the speaker or reference point. *M?e* is primarily endophoric, either anaphoric or cataphoric. Especially in its endophoric distribution, *m?e* is best viewed as ‘distance neutral’ (Himmelmann 1996: 211) since it can occur with all the CLFS to mark referents as ‘proximal/(coming) in the visual field’, ‘distal/(going) out of the visual field’, ‘never seen’, or depending on the particular classifier to refer to ‘horizontal’, ‘vertical’, or ‘bunched up’ referents. It is not accompanied by pointing gestures. It has also grammaticalised as a relativiser (§7.4).

The DEM2 set contains just the root *naqae* (Table 3). It takes CLFS as enclitics, unlike the DEM1 roots which take CLFS as proclitics. We analyse it as a ‘recognitional’ (RCG) demonstrative root, but in some contexts it may function more emphatically or mark unexpected information (§6).

Table 3: Pilagá recognitional demonstrative root (DEM2:RCG)

DEM2 root	Major sense	Instances in corpus
<i>Naqae</i>	‘this/that familiar but previously inactive; recognitional (RCG)’	84

Having now introduced the CLFS and demonstrative roots in Table 1–3, §3–§6 will discuss the morphosyntax and functions of four constructions that employ them. In anticipation, Table 4 surveys the grammatical functions of the basic classifier (BCLF), simple demonstrative (SDEM), deictic demonstrative (DDEM), and

recognitional demonstrative (RDEM) constructions. A dash in Table 4 indicates the morpheme in the first column lacks the adverbial function.

Table 4: Syntactic distribution of basic classifier (BCLF) and demonstrative constructions

	Adverbial	Pronominal	Determiner	Other
<i>ho?</i> DEM1 ‘proximal/ unspecified’	SDEM, DDEM	SDEM (rare), DDEM	DDEM	
<i>m?e</i> DEM1 ‘medial visible’; ‘neutral’	–	DDEM	DDEM	relativiser
<i>ča?a</i> DEM1 ‘distal visible’	–	DDEM	DDEM	
<i>maʃa</i> DEM1 ‘non-visible’	–	DDEM	DDEM	
<i>naqa(e)</i> DEM2 ‘recognitional’	–	DDEM	RDEM	
<i>da?</i> CLF ‘vertical; abstract’	–	BCLF	BCLF	subordinator
Other CLFS	–	BCLF (rare)	BCLF	

Across languages, demonstrative morphemes may have differing syntactic functions (Diessel 1999: 4; Krasnoukhova 2012). For example, in one language a single paradigm might function as demonstrative pronouns for participants or abstract concepts, as adnominal demonstratives, and as demonstrative adverbs for location or time. The Pilagá morpheme *ho?* distributes like this, though the particular construction it appears in (SDEM or DDEM) matters for syntactic function. In another language, a given demonstrative paradigm may have only a subset of functions. In English, for instance, *here/there* are adverbial demonstrative proforms

for locations,¹¹ and *now/then* are adverbial demonstrative proforms for time. But *this/that/these/those* function as both demonstrative participant pronouns and as demonstrative determiners.¹² The DEM1 roots and the DEM2 root distribute like these last English morphemes when in particular constructions. Classifiers in the BCLF construction function primarily as determiners, and more rarely as pronouns.

3 Basic classifier construction

In Pilagá discourse, determiners most frequently have the structure in (3). We call this the basic classifier construction (BCLF). The only required element is one of the six clitics in Table 1, or a plural counterpart. BCLFs functioning as determiners are highlighted in (1) above and in the examples below.

- (3) Basic classifier construction (BCLF)
(GENDER-)CLASSIFIER

The BCLF construction is illustrated in (4)–(6) with the posture/shape CLFs.

- (4) (028SanMartin2 1.5)
di? na?a-ik
CLF:HOR road-M
'road'
- (5) (190Verbos2 565)
da=m?e yi-la-?a da? epaq
CLF:VER=DEM1:NEUT A3-find-OBJ.SG CLF:VER tree
'She/He sees a tree.'
- (6) (028SanMartin2 1.2)
se-bide-w?o ñi? tamnaña-ki
A1-arrive-LOC:outward CLF:NO.EXT religion-place
'I arrive at the church.'

¹¹This sets aside uses like *this here dog*, where *here* doubles *this* as a proximal determiner.

¹²Diessel (1999: 90) also discusses presentational (what some call "predicational" or "identificational") and other functions of demonstratives.

Examples (7)–(8) illustrate vowel-lengthened plural CLFs. *Saa?* occurs in (8), rather than *soo?*, due to vowel harmony with the following noun. Recall that the plural CLF forms are optional (especially when the noun is marked for plurality).

- (7) **naa?** y-?ai?te
 CLF:near.PL POS1-eyes
 ‘my eyes’

- (8) (008ZorroPato 1)
 qanč?e yi-laey?a-lo **saa?** taʃaňi k'oqte-l
 then 3-see.ahead-PL CLF:far.PL rosy.billed.duck offspring-PL
 ‘He suddenly saw some rosy-billed ducklings.’

Examples (9)–(11) show the BCLF with gender prefixes. Masculine is usually unmarked (formally and functionally), but an overt prefix *ho-/(h)e-* can be added for clarity.

- (9) **ho-ga?** emek
 M-CLF:absent house
 ‘that (unknown) house’

- (10) (011Kitilipi 1.20)
 qač?e qo-i-law-lo **ho-?n** l-?ai?te ekey
 CONJ SBJ.INDF-A3-see-PL M-CLF:near POS3-eye.PL INTJ
 qo-d-?oya-lo **so?** l-?ai?te
 SBJ.INDF-A3-fear-PL CLF:far POS3-eye.PL
 ‘They saw the eyes (coming) and they got scared.’

- (11) (005ZorroAvispa 1.1)
 yi-la-?a **ha-so?** waʃat?o
 A3-find-OBJ.SG F-CLF:far wasp
 ‘They found a wasp (in the forest).’

Members of the deictic/visibility CLF subset in Table 1 can express metaphorical or cognitive distance, and sometimes a kind of evidentiality (§7).¹³ Example (12) describes customary actions. No particular mothers or carandillo palm leaves are

¹³Messineo et al. (2016) observe similar uses for the Toba cognates.

physically near the narrator, yet the ‘near’ CLF *na?* occurs. Example (13) is the first line of a folktale in which the participants are not departing from view within the world of discourse, though they are apparently conceptualised as distal and hence coded with the ‘far’ CLF *so?*

- (12) (039Artesania 1.1)

na? qad-at?e-l-pi da? set-ake d-?ofo-n-añan

CLF:near POS1PL-mother-PL-COL SUB want-DES A3-weave-NPROG-CAUS

nač?e w?ae-ñe yi-lake **na?** laqata l-awa

then be.first-COMPL A3-look.for CLF:near carandillo pos3-leaf

‘When our mothers want to weave (make handicrafts), first they look for carandillo (*trithrinax schizophylla*) leaves.’

- (13) (003Zorro Paloma 1.1)

w?o **so?** n-lo? so? waſayaql?ačiyi qataſa **so?**

EXIST CLF:far POS.INDF-day CLF:far fox and CLF:far

doqoto?

dove

‘There was a day the fox and the dove met each other.’

Similarly, (14) and (15) are the initial sentences of two different explanations of fishing customs. They seem to be situationally identical in objective deictic/visibility features, but in (14) generic ‘people’ who go fishing take the ‘far’ CLF, while in (15) generic ‘people’ who go fishing take the ‘near’ CLF *na?*. Presumably they are conceptualised differently within the world of discourse.

- (14) (013Pesca2 1.1)

so? siyaſa-di-pi da? set-ake di-y?ako

CLF:far person-PAUC-COL SUB want-DES A3-fish

‘When **the people** want to go to fish, ...’

- (15) (14Pesca4 1.1)

da? ni-y?akofo-k da? č?e n-piyae-yi da? di-y?ako **na?**

SUB B3-fish-M SUB suddenly B3-gather-PL SUB A3-fish CLF:near

siyaſa-di-pi

person-PAUC-COL

‘When it is fishing (time), when **the people** spontaneously gather to go hunter-gather (in general, lit. ‘fish’), ...’

The CLFs can also show psychological deixis in the sense of empathy or point-of-view. For instance, in (16) from a folktale, the skunk beats both the peccary (by killing the peccary with its odour and then eating it) and the fox (by outsmarting the fox). With one exception, the poor peccary is consistently referred to with the ‘near’ CLF in the story, while the skunk and the fox who eat or attempt to eat the peccaries are referred to with the ‘far’ CLF.

- (16) (004 ZorrinoZorro 1.4)

nač?e da? yi-lew na? owaqae, nač?e **so?** koñem ya-lik
 then SUB A3-die CLF:near peccary then CLF:far skunk A3-eat
ha-na=m?e owaqae
 F-CLF:near=DEM1:NEUT peccary

‘When the peccary dies, then the skunk eats this peccary.’

The third member of the deictic/visibility CLF subset is *ga?* ‘absent, unseen’. Its meaning ranges from ‘unseen now’ (i.e. absent, remote) to ‘never seen’ and hence ‘unknown’. Thus, it can indicate nonidentifiability or nonreferentiality, as in (17).¹⁴

- (17) (013Pesca2 1.1)

yi-k?atafa-som-?a **ga?** lačiya?ge
 A3-go-LOC:down-OBJ.SG CLF:absent stream

‘They (prepare to) go to a/some stream.’

Finally, a diminutive *t?ae(?)* can intervene between a CLF and a noun. As we will see below, this diminutive is becoming morphologised as part of demonstrative constructions.

- (18) (001ZorroPichi 2.10)

so? **t?ae** napam
 CLF:far DIM armadillo
 ‘the distant little armadillo’

We now turn to demonstrative constructions employing the roots in Table 2 and Table 3.

¹⁴ Also, some interrogative roots take the CLF *ga?* ‘unseen’, as in (22).

4 Simple demonstrative construction

The simple demonstrative (SDEM) construction contains only a demonstrative root and functions as an adverbial proform (cf. Table 4). This construction is limited to the ubiquitous DEM1 root *ho?*. Our understanding is that it is primarily used to draw the hearer's attention to something in the context, much as a pointing gesture does. In fact, the SDEM is often, but not always, accompanied by a physical gesture. As a simple demonstrative, *ho?* mostly functions as an exophoric adverbial locative, as in (19)–(21). *Ho?* is often translated as *aquí* ('here') but also as *allí* ('there') in Spanish. As an attention drawing form, it allows some locational range; but it is primarily proximal, so we gloss it consistently as 'proximal' to reflect this dominant use.¹⁵

- (19) (006ZorroCompanero 1.8)

a-w?afa-nyi	ho?	na?	yi-če
A2-HIT-LOC:MIDDLE DEM1:PROX CLF:NEAR POS1-LEG			
'Hit here (on) my leg!'			

- (20) (107Ethno26Grasa 3)

he-?n	četa	ho-ga-m?e	siyak	qanač?e
M-CLF:near grease M-CLF:absent-DEM1:NEUT animal then				
qo-y-ača-n-yi		ho?	ha-ga?	alewan?oſona
SBJ.INDF-A3-put-NPROG-LOC DEM1:PROX F-CLF:absent vessel				
'Then the fat of whatever animal they put there in an (earthenware) vessel (<i>alewan?oſona</i>).'				

- (21) (007ZorroWaqaw 1.8)

č?e	Ø-ek	ho?	de-mače-tape-get	so?=n-egaña-wa
soon A3-go DEM1:PROX A3-hear-PROG-VEN CLF:far=POS.INDF-friend-HUM				
waqa?w bird.species				

'Then, he (Fox) went away **there** [indicating the place where Waqaw was; not necessarily far or close], he heard his friend Waqaw (bird species) coming.'

Ho? can also have a temporal function, as in (22). (It also occurs in *ho(?)kal?io?* meaning 'before, long ago'.)

¹⁵ *He?n*, as in (20), is a common contraction from *he-na?*; the two forms are equivalent in meaning.

- (22) (001ZorroPichi 1.7)

qanč?e nae?=ga? aw-men ho? ñ-egaña-wa
 then INTG=CLF:absent A2-sell DEM1:PROX POS1-companion-HUM
 'So what will you sell **now**, my companion?'

Though the SDEM with *ho?* is primarily exophoric, it can be endophoric. In (23), it functions as a discourse anaphoric form, referring back to the act of being authorised to show a particular document.

- (23) (067ToribiaAcosta.46–48)

- a. hayem ka? sepa č?e algún documento da? da?
 1SG before seem.to.me then some document SUB SUB
 Ø-ek-a so? saλa-nek
 3-go-LOC:specific CLF:far chief-AGENT
 'I believed that, that the chief came back with the document'
- b. da? qomi? y-aloflo-na-lo
 SUB 1PL 3-show-NPROG-PL
 'so that he could show us.'
- c. da? ka? epaña autorisaw ho? eta-t
 SUB before it.seems authorised DEM1:PROX say-PROG
 'He was saying that he seems authorised for **this**'

In the more complex demonstrative construction next discussed in §5, we find *ho?* in both adverbial and non-adverbial functions.

5 Deictic demonstrative construction

Pilagá has a complex deictic demonstrative (DDEM) construction involving the DEM1 roots (Table 2) plus the CLFS (Table 1). The elements of complex demonstratives show dialect and idiolect variation and may vary by speaker's age. As noted in §2, some elements can undergo vowel harmony. The 'near' CLF *na?* often reduces to (?)*n*, and the 'neutral' DEM1 root *m?e* often reduces to (?)*m*. There is considerable variation in the text corpus especially for *m?e*. For instance, *hogam?e*, *hogam?o?*, and *hogan?e* all contain *m?e* and are alternative forms of 'that absent/unknown'. *Ho?n* is a contraction from *ho-na?=m?e* (M-CLF:near=DEM1:NEUT). According to the consultant Ignacio Silva, some of the variant forms are "old words", rarely heard now. All these factors result in a great variety of surface forms.

The deictic demonstrative (DDEM) construction has the structure in (24).

- (24) Deictic demonstrative construction (DDEM)
 (GENDER-)CLASSIFIER=(diminutive=)DEM1(-plural)

The gender markers in the DDEM are *ha-* ‘feminine’ and *(h)e-/ho-/Ø* ‘masculine’, illustrated in (25)–(28). Sometimes the masculine is left unmarked for gender. We do not write the zero form in examples. Plural can be marked by lengthening the CLF vowel, and some DDEMs add *-lo* or *-wa* ‘plural’.

- (25) (004ZorrinoZorro 1.4)
 ha-na=m?e owaqae
 F-CLF:near=DEM1:NEUT peccary
 ‘this peccary’

- (26) (136ethnograph55 6)
 naqae=ga ho-ga=ma?f piy?o?onaq
 DEM2:RCG=CLF:absent M-CLF:absent=DEM1:NVIS shaman
 ‘(Death could result from the action of) **some/any** shaman.’

- (27) (001ZorroPichi 2.14)
 yeči ki hora da? ho-da=ma?f y-em
 evident what hour SUB M-CLF:VER=DEM1:NVIS A3-end
 ‘(I don’t know) what time **that** (the story) ends...’

- (28) (032ColoniaEnsanchez 1.3)
 ñi=ma?f ñi? qan-saλa?f-nek
 CLF:NO.EXT=DEM1:NVIS CLF:NO.EXT POS1PL-chief-M
 ‘**that** our chief’ (not present at the time of utterance)¹⁶

‘Shape’ (rounded) or ‘size’ appears as a semantic extension of ‘feminine’ gender. However, not all nominals in Pilagá are marked for a particular gender distinction, regardless of their shape, nor is such marking synchronically predictable. As (24) indicates, a diminutive can occur between the CLF and demonstrative root, as in (29)–(30). The diminutive is acceptable after a DEM1 root only

¹⁶ A native speaker said this text line sounded redundant, apparently due to both the DDEM and the separate BCLF before ‘our chief’.

if the diminutive is preceded by a CLF (as if the diminutive morpheme is nominal); compare (31)–(32). The diminutive can communicate that one is feeling sorry for a referent.

- (29) *ñi?*=*t?ae*=*m?e*

CLF:NO.EXT=DIM=DEM1:NEUT

‘that little rounded/sitting one’ (I may be seeing it or not)

- (30) *da?*=*t?ae*=*ča?**a*

CLF:VER=DIM=DEM1:DIST.VIS

‘that far little one’ (I see it)

- (31) *ñi?*=*m?e* *ñi?* *t?ae*

CLF:NO.EXT=DEM1:NEUT CLF:NO.EXT DIM

‘that little rounded/sitting one’ (I may be seeing it or not)

- (32) **da?*=*ča?**a* *t?ae*

CLF:VER=DEM1:DIST.VIS DIM

All members of the DEM1 paradigm (Table 2) occur in the DDEM construction. We illustrate this in combination with the ‘horizontal’ CLF *di?*. In (33), *ho?* indicates the object is close to the speaker and visible at the time of utterance. *M?e* is neutral in (34) about whether the object is visible at speech time. *Ča?**a* in (35) requires that the object be visible at speech time. *Ma?**a* in (36) indicates the object is not present/visible to the speaker at speech time.

- (33) *yi-la?**a* **di=ho?** *siya?**a**wawa*

A3-see CLF:HOR=DEM1:PROX person

‘She/He saw **this** person lying down/asleep/dead.’ (The person is visible now and close; but need not currently be horizontal/dead.)

- (34) *yi-la?**a* **di=m?e** *siya?**a**wawa*

A3-see CLF:HOR=DEM1:NEUT person

‘She/He saw **a/that** person lying/sleeping/dead.’ (The person may or may not be in sight at the time of speaking.)

- (35) yi-la? a di=ča? a siya?awa
 A3-see CLF:HOR=DEM1:DIST.VIS person
 ‘She/He saw that far-away lying-down/asleep/dead person.’ (The person is visible now and far away; pointing to the person.)
- (36) yi-la? a di=ma? a siya?awa
 A3-see CLF:HOR=DEM1:NVIS person
 ‘She/He saw that person lying down/asleep/dead.’ (The person is not visible to the speaker.)

Though all DEM1 roots occur in the DDEM construction, there are some co-occurrence restrictions with particular CLFS to avoid semantic clashes. This is particularly relevant for the deictic/visibility CLFS, as the posture/shape CLFS do not lend deictic information to the overall meaning of the demonstrative construction (as seen in (33)–(36) with *di?* ‘horizontal’).

Ča? a ‘distant visible’ only occurs with CLFS that allow interpretation of a visible referent, i.e. *na?* ‘near, coming’, *so?* ‘far, departing’, and the posture/shape CLFS, as in (37)–(41). Examples (37)–(38) have a distal+visible referent, marked by ča? a. The fact that it is approaching the reference point (potentially communicated by *na?*) may be communicated with or without the ‘ventive’ suffix -get on ča? a. The ‘itive’ -ge(?) is not possible with *na?=ča? a*, but the ‘itive’ is possible with *so?=ča? a*, as in (39).

- (37) na?=č? a
 CLF:near=DEM1:DIST.VIS
 ‘far referent coming near’
- (38) na?=č? a-get
 CLF:near=DEM1:DIST.VIS-VEN
 ‘far referent coming near’
- (39) so?=ča? a-ge?
 CLF:far=DEM1:DIST.VIS-IT
 ‘far referent going away’
- (40) do=ča? a
 CLF:VER=DEM1:DIST.VIS
 ‘that upright far referent’

- (41) ñia=č?a-lo
 CLF:NO.EXT=DEM1:DIST.VIS-PL
 'those sitting there'

The three examples sets between (42) and (52) illustrate additional combinations of the deictic/visibility CLFs with the more frequent DEM1 roots. The specific interpretation of a combination may depend on pragmatic context. Examples (42)–(44) carry the ‘near, coming’ CLF *na?*. In (44), the ‘distal’ feature of the DEM1 root *ča?a* overrides any ‘near, proximal, coming’ meaning that might otherwise be associated with *na?*; this suggests that *na?* may be bleaching of its spatial semantics. Along with a wave of the hand, (44) could serve as an answer to the question ‘Where is José?’

- (42) **no?=ho?** na? t?ae
 CLF:near=DEM1:PROX CLF:near DIM
 ‘this little one’ (right here beside me and I see it)

- (43) **na?=m?e** na? t?ae
 CLF:near=DEM1:NEUT CLF:near DIM
 ‘this little one’ (the item may be present or not; the expression could refer to something I have been talking about)

- (44) **na?=ča?a** na? t?ae
 CLF:near=DEM1:DIST.VIS CLF:near DIM
 ‘(he’s) that little one’ (there, not moving)

Examples (45)–(49) combine *so?* with DEM1 roots. Our consultant found (45) unacceptable, explaining that it contradictorily combines *so?* ‘far’ with *ho?* ‘proximal’ (we return to this combination further below). When *so?* combines with ‘neutral’ *m?e*, as in (46), the result indicates a visible or identifiable referent departing from the deictic center; thus with *m?e*, the CLF yields the primary deixis/visibility meaning. In (47) with *ča?a* ‘distal visible’ plus the ‘itive’ -*ge*(?), the overall reading is of an already distal but visible participant moving away. Without the ‘itive’ (48), one consultant finds *so?* incompatible with *ča?a*. This is because *so?* can sometimes be interpreted as ‘(going) out of view’, while *ča?a* specifically indicates ‘visible’; but the combination was acceptable in (39). In (49) with *mafa* ‘non-visible’, the speaker could possibly know the non-visible referent, though there is something uncertain about it in the speaker’s mind.

- (45) *soʔ=tʔae=hoʔ

CLF:far=DIM=DEM1:PROX

- (46) soʔ=tʔae=mʔe

CLF:far=DIM=DEM1:NEUT

‘that small visible/identifiable departing referent’

- (47) soʔ=tʔae=čaʔa-ge

CLF:far=DIM=DEM1:DIST.VIS-IT

‘that small visible far-away departing referent’

- (48) *soʔ=tʔae=čaʔa

CLF:far=DIM=DEM1:DIST.VIS

- (49) soʔ=tʔae=maɻa

CLF:far=DIM=DEM1:NVIS

‘that little (stationary) unseen referent’ (perhaps I know it)

Possible interpretations of the ‘absent’ CLF *ga?* include unknown, non-specific, and non-referential readings, as in (50)–(53). It may combine with the ‘proximal’ and ‘neutral’ DEM1 roots, but not with *čaʔa* ‘distal+visible’, as shown by (50)–(52). This restriction is due to the semantic clash between the ‘visible’ feature of *čaʔa* and the ‘absent’ feature of *ga?*. The perhaps surprising example in this set is (50), as it might seem that the ‘absent’ feature of *ga?* should conflict with *ho?*. However, its acceptability reveals the expanding semantic domain of polysemous *ho?*; in particular, with a CLF, *ho?* can be used endophorically for a participant currently under discussion. In this use, it participates in topic marking. Example (51) shows that with the ‘neutral’ DEM1 root, the semantic features of the CLF again become especially evident.

- (50) gaʔ=tʔae=hoʔ

CLF:absent=DIM=DEM1:PROX

‘this little one’ (referring to something/somebody under discussion that is far or I do not remember well)

- (51) gaʔ=tʔae=mʔe

CLF:absent=DIM=DEM1:NEUT

‘that little one’ (not in view, never seen, or unknown, but I remember it)

- (52) *gaʔ=taʔe=čaʔa
 CLF:absent=DIM=DEM1:DIST.VIS

If under the scope of negation, *gaʔ* plus *mʔe* may indicate ‘nothing, nobody’, as in (53).

- (53) Qaya gaʔ=mʔe
 NEXIST.HUM CLF:absent=DEM1:NEUT
 ‘There is **nobody**.’

In (44), we saw that the meaning of DEM1 čaʔa overrides the spatial meaning that CLF *na?* might otherwise carry. However, in some situations the meaning of a CLF can override that of a DEM1 root. Thus, (54)–(57) were said to mean “basically the same” in terms of spatial/visibility deixis. They all carry the CLF *so?* ‘far, departing (potentially to the point of being absent)’, regardless of choice of the demonstrative root. Notably, (55) was judged as fine, while (45) with the same key elements was rejected. We analyse the variability in speakers’ judgments as reflecting the polysemous nature of *ho?*: on one occasion its exophoric ‘proximal’ feature is conceptually prominent and thus it is viewed as conflicting with *so?*, but on another – as in (55) – *ho?* is interpreted endophorically to indicate the participant under discussion in the discourse, so there is no spatial deixis conflict. The DEM2 root in (57) is discussed in §6.

- (54) so? y-alek
 CLF:far POS1-son
 ‘my son (distant/absent)’

- (55) so=ho? y-alek
 CLF:far=DEM1:PROX POS1-son
 ‘that my son (departing)’

- (56) so=mʔe y-alek
 CLF:far=DEM1:NEUT POS1-son
 ‘that my son (distant/absent)’

- (57) naqae=so? y-alek
 DEM2:RCG=CLF:far POS1-son
 ‘that (is) my son’ (understood to not be present)

The preceding discussion has focused on structure of the DDEM and meanings of composing morphemes. We now more explicitly address grammatical and discourse functions of this construction (cf. Table 4). DDEM constructions serve as adverbial and participant proforms or as determiners. The proform function is illustrated in (58)–(60) with the root *ho?*. The two senses of (58) show the adverbial exophoric locative function of the DDEM with *ho?*, and its participant reference function. In (59), the DDEM refers exophorically to an inanimate entity. In (60), it refers exophorically to a location.

- (58) **so=ho?**

CLF:far=DEM1:PROX

‘there’ (Spanish *allá*) / ‘one (who is) departing’

- (59) **ha-n=ho?** mate

F-CLF:near=DEM1:PROX mate(drink)

‘This is a mate (container).’

- (60) (060TrabajoMadera 12)

maλaña qaga ha-ñó=ho? naña na-ñ-?a
yet NEXIST F-CLF:NO.EXT=DEM1:PROX now B3-sit-OBJ.SG
dyo=ho? Campo
CLF:HOR=DEM1:PROX place.name

‘This [pointing to the location of the community] did not yet exist (which) is now (the spread-out community of) Estanislao del Campo.’

Example (59) is a zero-copula equational clause and the DDEM is not in the same phrase as *mate*. The pronominal DDEM with *na?* ‘near’ plus *ho?* indicates an item close enough to touch. The feminine gender prefix occurs due to the rounded shape of the container. In (60), *ha-ñó=ho?* indicates a non-extended referent, pointing to the particular location (rather than extended shape) of the community.

Deictic demonstrative constructions with DEM1 roots other than *ho?* function only pronominally and as determiners (not adverbially; cf. Table 4). In (61), *he-n=m?e* functions as a text-internal anaphoric pronominal. It refers to the story the speaker is in the midst of relating. (*He-n=ho?* with the ‘proximal’ DEM1 root in this context would mean ‘here, the place where I, the speaker, am’.)

- (61) (001ZorroPichi 2.13)

he-n=m?e huw!
 M-CLF:near=DEM1:NEUT WOW
 ‘this (story), wow!’ (meaning ‘this story I am telling you’)

In (62), the highlighted DDEM functions as a cataphoric pronominal. *M?e* carries the ‘vertical’ CLF *da?*, but in this context it designates a propositional event which, as a whole, is an abstract concept.

- (62) (048RecoleccionMiel 1.1)

- a. so=m?e siya?fa-di-pi da?=m?e
 CLF:far=DEM1:NEUT person-PAUC-COL CLF:VER=DEM1:NEUT
 qo-ila-?a so? kon'aya?apo?o?
 SBJ.INDF-see-OBJ.SG CLF:far bee.hive
 ‘When the men find the bee hive,
- b. nač'e w'ae-ñe da?=m?e
 soon be.first-COMPL CLF:VER=DEM1:NEUT
 ‘they first do this:’
- c. qo-ya-lo-n so? dole?
 SBJ.INDF-A3-stir-NPROG CLF:far fire
 ‘they stir up the fire.’

A DDEM with *ča?a* may function exophorically or endophorically. The exophoric function is dominant, but in (63d), from a story about competition between Fox and Toad, *ni?=ča?a* is endophoric, referring to the toad. Line (63d) also shows the pronominal DDEM *so?=t?ae=m?e* functioning anaphorically.

- (63) (002SapoZorro 1.11–1.14)

- a. degesesow eso wayqal'ačiyi yači ena?aye-ik
 quickly CLF:far fox certain dusty-AUG
 ‘Quickly, it is clear/certain that the fox stirred up a lot of dust.’
- b. so? qololo da? Ø-wenot qanač?e yita?fa ne-no?o-segem so?
 CLF:far toad SUB A3-jump then again B3-move-upward CLF:far
 qololo l-qaya
 toad pos3-sibling
 ‘But when the toad jumped, another toad appeared.’

- c. ta Ø-wenot ta ne-noñó-segem so? qololo l-qaya
 AGAIN A3-jump again B3-move-upward CLF:far toad POS3-sibling
 ye-d?̄-a-ta
 A3-arrive-OBJ.SG-other.side
 ‘He jumped again, and another toad (appeared until) it reached the finish.’
- d. qanč'e so?̄=t?ae=m?e qanč'e yeči, n-selka-pe-get
 then CLF:far=DIM=DEM1:NEUT then certain B3-see-PROG-VEN
 ñi?̄=ča?a=w ñi?̄ n-qomit-aña-wa
 CLF:NO.EXT=DEM1:DIST.VIS=INTSF CLF:NO.EXT B3-compete-NMLZ-HUM
 ñi?̄ yači yi-we?en
 CLF:NO.EXT certain A3-laugh
 ‘He (Fox) certainly keeps looking for that far-distant one (toad)
 coming towards him and so he (toad, referenced throughout by *ñi?*)
 certainly/evidently laughs at the competitor (Fox).’

Examples (64)–(65) show pronominal DDEMS with the ‘nonvisible’ root *masa*.

- (64) (032ColoniaEnsanchez 1.1)
- di?̄=maña di?̄-ae qad-?ačaqa? le-nañat Colonia
 CLF:HOR=DEM1:NVIS CLF:HOR-F POS1PL-community POS3-name Colonia
 Ensanchez Ø-na?a-ge na? se?w
 Ensanchez A3-be-it CLF:near north
 ‘This our community, its name (is) Colonia Ensanchez, is towards the north.’ (Context: The speaker is in a workshop talking about his far-distant community, probably looking at a map.)

- (65) (032ColoniaEnsanchez 1.3)
- ñi?̄=maña ñi?̄ qan-saλaña-nek l-sek
 CLF:NO.EXT=DEM1:NVIS CLF:NO.EXT POS1PL-chief-M POS3-neighbor
 ha-ñi?̄ tamnaña-ki
 F-CLF:NO.EXT religion-LOC
 ‘That (far away house) (that I’m talking about from memory) of our chief is between the church (and the school).’

We now briefly comment on adnominal DDEM uses. In (66), *dyo=ho?* refers exophorically to a concrete participant. In (67), *he-n=ho?* refers to a time.

- (66) **dyo=ho?** pioq če?eda weta-ñ-?a kali?o
 CLF:HOR-DEM1:PROX dog be.first be-LOC:below-OBJ.SG long.ago
 ‘This dog (present, that I am signaling) has been (lying) here a long time.’

- (67) (077Sent09Cantidad 2)
 so? l-aqaya setae? na-paſagen-a da? paſagentanaſai
 CLF:far POS2-brother want B3-learn-OBJ.SG SUB teacher.M
he-n=ho? wo?e
 M-CLF:near=DEM1:PROX year
 ‘Your brother wants to study teaching (to be a teacher) **this** year.’

Adnominal DDEMs with *m?e* often mark already-mentioned participants, as in (16). But this is not always the case. In (68), *di?=m?e* occurs on the first mention of ‘garden/field’; the consultant expressed the view that the sentence would mean essentially the same thing if a BCLF with just *di?* occurred instead.

- (68) (035Linea 1.4)
 da? set-ake a-e-ye **di?=m?e** qad-an-aſan-qa?
 SUB want-DES A3-go-in.line CLF:HOR=DEM1:NEUT POS1PL-plant-CAUS-LOC
 qanač'e o-ket'a-ge di? na?a-ik Ø-leka?a-ege
 then A2-continue-IT CLF:HOR road-M 3A-be.big-forward
 ‘If you want to get to our vegetable garden, you have to continue along the wide path.’

In (69), *m?e* combines with the ‘absent’ CLF *ga?*, to determine the nonreferential phrase ‘our thought’

- (69) (001ZorroPichi 1.1)
 čaqaga da? **ga?=m?e** qad-enat-aſak
 what SUB CLF:absent=DEM1:NEUT POS1PL-think-NMLZ
 ‘What (is) our thought? (i.e. ‘What shall we do?’)

6 Recognitional demonstrative construction

A second demonstrative construction has not, to our knowledge, been noted in previous Guaykuruan literature. We call this a recognitional demonstrative (RDEM) construction. In it, the root *naqae* co-occurs with a CLF, but *naqae* differs from the DEM1 set in taking the CLF as an enclitic, yielding the structure in (70).

- (70) Recognitional demonstrative construction (RDEM)
DEM2.root=CLASSIFIER(-plural)

Naqae indicates that the speaker anticipates the hearer already knows or is familiar with the identity of the referent (whether or not it has already been mentioned in the discourse), but wishes to activate it in the hearer's mind. There may be an assumption of shared knowledge about the referent, but there may be doubt or even disbelief that the hearer is currently attending to it, so the speaker is activating it for the hearer. This is similar to what Himmelmann (1996) and Diessel (1999) call a "recognitional" demonstrative. We consider *naqae* to be a demonstrative root as it orients the hearer's attention to a participant.

Though consultants specifically comment that *naqae* means the hearer knows the referent, in some contexts we think *naqae* would be better characterized as indicating a familiar concept, as it can also be used for non-referential mentions. Speakers suggest it sometimes indicates a note of surprise or unexpectedness about a known but previously inactive referent, as if something has just activated it in the mind of the speaker. This is the case in (71), which stacks *naqae=ñi* together with *ho?* and *ñi?*. Here, *ho?* is verbally signalling (verbally "pointing") to the person, who is sitting.

- (71) naqae=ñi ho? ñi? siyaʃawa
DEM2:RCG=CLF:NO.EXT DEM1:PROX CLF:NO.EXT person
'Ah, that/this is the person!' (I see him/her, sitting)

The RDEM construction is attested in pronominal and adnominal functions (cf. Table 4). In (72), *naqae=na-wa* functions pronominally.¹⁷ *Naqa=ñi* is also pronominal in (73). However, *naqae=na?* in (73) appears to be adnominal. In the discourse just prior to (73), the fox is annoyed by a wasp and says, "Why are you always in my path? I'm going to hit you." Fox then utters (73). Here, *naqaena?* indicates some emotiveness or unexpectedness.

- (72) (017Pesca1 1.4)
qataʃa da? an-awa-?n naa? l-?awaʃak-o
and SUB B2-watch-PL-NPROG CLF:near.PL POS3-water.channel-PL
naqae=na-wa naa? n-aya-pe-eg?a-lo
DEM2:RCG=CLF:near-PL CLF:near.PL B3-leave-PROG-LOC:specific-PL

¹⁷The second instance of the CLF *naa?* in (72) functions like a relativiser to introduce a clause modifying *naqae=na-wa*.

he-n ñiyaq-pi

M-CLF:near fish-COL

'Also to watch the water channels, **these** are where from the fish emerge.'

- (73) (005ZorroAvispa 1.2)

l?e? naqa=ñi y-?ata-?nyi naqae=na?

why DEM2:RCG=CLF:NOEXT A3-move-LOC:middle DEM2:RCG=CLF:near

y-adik

POS1-path

'Why does **this one** move (be) in **this** my path?'

The RDEM construction can be anaphoric. In (74d), *naqae=na-wa* refers back to 'the place where the fish pass' mentioned in (74b).

- (74) (017Pesca1 1.1–104)

a. w'ae-ñi qomi qo-ya-paʃage-nek-e da?

be.first-COMPL 1PL SBJ.INDF-A3-teach-AGENT-PL SUB

qo-y-e?et na? čikena

SBJ.INDF-A3-prepare CLF:near arrow

'First they taught us to prepare the arrows'

b. qataʃa na? Ø-wapiň'a-lo qataʃa na?

and CLF:near 3-be.place-PL and CLF:near

n-aeya-pe-ege-?a na? ñiyaqa-pi

B3-go-IPFV-opposite-OBJ.SG CLF:near fish-COL

'and (to know) the places and where the fish pass by'

c. qataʃa qomi qo-ya-paʃage-nek-e da?

and 1PL SBJ.INDF-A3-teach-AGENT-PL SUB

qo-ya-ye-n na? ñiyaq

SBJ.INDF-A3-throw-NPROG CLF:near fish

'Also they taught us how to stab a fish'

d. qataʃa da? an-awa-?n naa? l-?awaʃako

and SUB B2-watch-pl-nprog CLF:near.PL POS3-caudal

naqae=na-wa naa? n-aya-p-ege-lo

DEM2:RCG=CLF:near-PL CLF:near.PL B3-go-PROG-opposite-PL

he-?n ñiyaqa-pi

M-CLF:near fish-COL

'and how to watch the flow where the fish come out.'

Finally, (26) suggests that Pilagá allows stacking of RDEM and DDEM.

7 Further grammaticalisation: Nominal TAM and subordination

7.1 Overview

Having now discussed the morphosyntax and basic functions of CLFS and demonstratives, we turn to extended uses for nominal tense, mood/evidentiality, and clausal subordination. Pilagá adds to the body of data showing how demonstratives and determiners can further grammaticalise (Diessel 1999; 2003; Gildea 1993; Aikhenvald 2015).

7.2 Incipient nominal tense, mood, and evidentiality

Like other Guaykuruan languages, Pilagá lacks grammatical tense forms. However, some CLFS and DEM1 roots implicate temporal meanings in certain contexts, and visible versus inferred source of evidence or (un)certainty. The temporal and evidentiality/modality meanings sometimes relate to evaluation of a nominal referent and sometimes to the proposition. Pilagá thus pertains to the set of languages having nominal TAM (Nordlinger & Sadler 2004). The role of CLFS in conveying temporal, modal, and evidential meanings in Guaykuruan has been discussed in other works (Messineo et al. 2016; Messineo & Cúneo 2019), including for Pilagá (Vidal & Klein 1998; Vidal & Gutiérrez 2010). Here we also note the role of DEM1 roots in marking these concepts.

In Pilagá, temporal use of CLFS and DEM1 roots is pragmatic rather than fully grammaticalised, and interpretations interact with person and lexical meanings. First, (75)–(76) reveal the possible present-time interpretation of posture/shape CLFS versus the past-time effect of *so?* ‘far, departing’. *Da?* is the CLF for abstract nouns like *lasook* ‘custom’, as well as for vertical ‘person’. The overall interpretation in (75) is present time. In (76), *so?* occurs with both nouns. Given the abstract concept of ‘custom’, *so?* cannot be interpreted as meaning that *lasook* is spatially distant or moving away, so a space-to-time metaphorical inference yields the understanding of a ‘distant’ or past time situation. This likely also affects the use and interpretation of *so?* with ‘person’.

- (75) eta ho? **da?** lasook da? siyaʃawa
it.is.said DEM1:PROX CLF:VER custom CLF:VER person
‘This is the custom of the person.’ (present)

- (76) eta ho? so? lasook so? siyaña-di-pi
 it.is.said DEM1:PROX CLF:far custom CLF:far person-PAUC-COL
 ‘This was the custom (of) the people.’

To more clearly see the possible temporal effect of *so?* when applied to concrete objects, consider (77). *Ho?* occurs in *so?=ho?* because ‘my son’ is in the speaker’s vicinity at the time of utterance. Since ‘my son’ is locally present, *so?* ‘far’ can only be interpreted as indicating a temporally distant or past event. In this instance the CLF has propositional/event-scope, while the DEM1 root has nominal scope related to the speech time.

- (77) (052RelatoAnciana 62)
 n-oye-tak naña so?=ho? y-alek n-woñom da?
 B3-cry-PROG now CLF:far=DEM1:PROX POS1-son B3-feel CLF:VER
 l-qowaña
 pos3-hunger
 ‘My son here/now was crying because he felt hunger.’

So? does not obligate a past-time propositional interpretation if contextual factors indicate otherwise. Because of *qomle* ‘later’ in (78), *so?* is interpreted as applying to the past-time of the events involving ‘our ancestors’ and not to the event of telling.

- (78) (Vidal & Gutiérrez 2010: 1353)
 qomle s-aqtanañan so? qadetalpi
 later A1-tell CLF:far our.grandparents
 ‘I’m going to tell you about our ancestors.’

In contrast to *so?*, the CLF *ga?* ‘unseen’ pragmatically allows that the “event in which it is embedded is an expression of the ignorance, the desires, or the intentions of the speaker, rather than a realized event” (Vidal & Klein 1998: 176). *Ga?* often occurs in clauses with conditional, obligation, or prospective meaning, as in (79).

- (79) (025EspirituSuri 1.3)
 awa-w?o-e ga? ade-wo ...
 A2-make-PL CLF:absent POS2-clothes
 ‘you have to make your costumes ...’

Temporal interpretation is affected by pragmatic interaction between person, proximity of a referent to the speaker versus other referents, and the semantics

of lexemes, CLFs and DEM roots. In (80)–(82), the speaker and the grammatical subject are the same person. *Taqa* ‘talk’ plus *ñó=ho?* ‘non.extended=proximal’ implies a present-time action because the first-person speaker can talk ‘now’ to someone who is physically near.

- (80) se-taqa-tap-ege *ñó=ho?* siyañawa
 A1-talk-PROG-IT CLF:NO.EXT=DEM1:PROX person
 ‘I am talking to a/this person (sitting next/close to me).’

In (81) with ‘talk’, the ‘neutral’ DEM1 root *m?e* with a posture/shape CLF allows a present or past interpretation. In (82), ‘distal+visible’ *ča?á* implies a past event because – ignoring telephones – one cannot talk ‘now’ to someone far away.

- (81) se-taqa-tap-ege *ñí=m?e* siyañawa
 A1-talk-PROG-IT CLF:NO.EXT=DEM1:NEUT person
 ‘I am talking now to a/that person (who is sitting).’ /
 ‘I talked to that person (who is now sitting).’

- (82) se-taqa-tap-ege *ñí=ča?á* siyañawa
 A1-talk-PROG-IT CLF:NO.EXT=DEM1:DIST.VIS person
 ‘I was talking to that person now sitting (far from me).’ (Since he/she is far away, it is impossible to be talking to him/her right now.)

DEM1 roots also play a role in expressing a speaker’s (un)certainty. Compare (83)–(84), which show that *maña* is a marker of uncertainty compared to *m?e*.

- (83) eta ho-da=*maña* lasook
 it.is.said M-CLF:VER=DEM1:NVIS custom
 ‘It is said that is how the custom must have been.’ /
 ‘That seems to have been (how) the custom (was).’

- (84) eta ho-da=*m?e* lasook
 it.is.said M-CLF:VER=DEM1:NEUT custom
 ‘It is said this is what the custom is (like).’ (speaker is certain)

7.3 *Da?* as clausal subordinator

Elements of the determiner and demonstrative systems have become markers of subordination (Vidal 2001). The CLF *da?* ‘vertically extended’ introduces clauses

with a variety of adverbial, complement, and nominal-modifying functions. This needs more exposition than can be taken up here, but we note that it introduces readings of at least adverbial ‘when’ in (15)–(16), ‘conditional’ in (68), and ‘purpose’ in (85). In (77), *da?* occurs before an abstract nominal ‘hunger’, but the phrase with *da?* communicates an adverbial ‘because’ notion. The complement function is illustrated in (23c), (74c), and (86), and a nominal-modifying (i.e. relative) function surfaces in (27) and (87).

- (85) (004ZorrinoZorro 1.1)

so?	koñem	w?o	so?	maeč?e	la-wa-nafanqa?	da?
CLF:FAR	skunk	exist	CLF:far	own	POS3-trap-NMLZ:place	SUB
na-wa-n	na?	owaqae				
B3-trap-NPROG	CLF:near	peccary				

‘The skunk had his own trapping place **in order to** trap the peccary.’

- (86) (015Pesca3 1.2)

wač?e	qo-d-?oya	da?	ne-matae-yi	ga=m?
because	SBJ.INDF-A3-fear	SUB	B3-puncture-PL	CLF:absent=DEM1:NEUT
n-ofonek				

POS.INDF-fish

‘Because they feared **that** the fish would damage it.’

- (87) (003ZorroPaloma 1.2)

yi-pit-etpa-lo	sa-wa	l?-ai?te	so?	doqoto?	da?	tořomaqčiglo
A3-want-PROG-PL	CLF:far-PL	POS3-eyes	CLF:far	dove	SUB	be.red

‘He wanted dove eyes **that** were red.’

Historically, the *da?* subordinator is likely connected to the ‘abstract’ nominal determining function of *da?*. As *da?* is the CLF for abstract nominal referents, it is well-suited to mark nominalised propositions, which are typically rather abstract conceptual entities. These then come to serve as subordinate clauses.

7.4 *Mʔe* as a relativiser

Mʔe is a highly frequent demonstrative root (Table 2). We have seen that in contrastive elicitation, it allows a ‘medial’ spatial contrast between *ho?* ‘proximal’ and *čaʔa* ‘distal visible’, and a visibility contrast with *maša* ‘not visible’. However, it can occur with all deictic CLFs to mark referents as ‘proximal/in the visual

field', 'distal/(going) out of the visual field', or 'never seen/absent/nonreferential'; and it occurs with all posture/shape CLFS. We also noted that *m?e* demonstratives can be used cataphorically, as in (62), though they are usually anaphoric. Given its range of collocations and uses, we conclude that *m?e* has developed a 'distance/deictically neutral' role (Himmelmann 1996: 211).

Perhaps concomitant with its neutral deictic use, *m?e* has developed as the most common relativiser. It follows a head noun to anaphorically introduce a modifying relative clause, as in (88)–(89). As a relativiser, it does not occur with classifiers or gender affixes (this is also true of the Western Toba cognate; Carpio 2012: 53). It thus diverges from the DDEM construction involving this root, which requires a classifier.

- (88) (052RelatoAnciana 52)
- | | | | |
|--|-------------|-------------------|------------------|
| ad-apena? | l-t?a | di?=m | ad-apena? |
| POS2-grandfather | POS3-father | CLF:HOR=DEM1:NEUT | POS2-grandfather |
| m?e | yi-wa | | |
| DEM1:NEUT | POS1-spouse | | |
| ‘the father of your grandfather (deceased) that was (my) husband’ | | | |

- (89) (071Sent03Comunidad 7)
- | | | | | | |
|---|-------------|------|-----------------|------------|---------------|
| naega? | wa?a-ege | nqo? | ga?=nadik | m?e | yi-lot-?a |
| where | be-opposite | when | CLF:absent=road | DEM1:NEUT | A3-see-OBJ.SG |
| ga?=Joel | | | | | |
| CLF:absent=Joel | | | | | |
| ‘Where is the road that goes directly to (lit. sees) (the house of) Joel?’ | | | | | |

M?e also introduces headless relative clauses, as in (90).

- (90) (013Pesca2 1.4)
- | | | | | |
|---|-------|-----------------|-----------|---------------------|
| yi-la?a-ge | nač?e | yi-lo?t-ege | m?e | t-a-y-?a |
| A3-see-IT | soon | A3-see-opposite | DEM1:NEUT | A3-go-inside-OBJ.SG |
| ‘He follows it (a bee, with his gaze) to see (the place) where it goes inside (of the honeycomb).’ | | | | |

The relativising use of *m?e* might at first appear to be the SDEM construction; but by itself, *m?e* is not synchronically attested as a proform. Nevertheless, it is largely associated with discourse anaphoricity. It has become the unmarked means to refer to a just-mentioned referent. Historically, this may have come about via an adjoined clause. That is, a conceivable earlier analysis of (89) might

have been ‘Where is the road, **that one** (i.e. ‘road’) sees Joel?’ The relativiser function then developed by reanalysing the modifying clause (‘that [one] sees Joel’) as embedded. If this scenario is correct, then contra Himmelmann’s (1996) suggestion, it is not the distal member of the demonstrative paradigm which has extended its meaning to become grammaticalised as a relativiser, but the ‘middle (visible)’ and/or ‘neutral’ member of the paradigm.

8 Conclusions

This study contributes to our understanding of the typological range of determiner and demonstrative systems. It has especially highlighted the demonstrative roots, which have not received much previous study in Guaykuruan languages.

Anyone who has examined a substantive discourse sample for any language, and over that sample tried to specify “all and only” the componential semantic features that distinct demonstrative forms have, can surely attest that choice among demonstrative morphemes cannot be tied exclusively to literal spatial deixis nor to “clean” endophoric versus exophoric factors. The choice is always sensitive to the speaker’s conceptualisation of referents on particular occasions of speaking, and to assumptions about the hearer’s continually changing state of mind in the endeavour to establish joint attention. With these important cautions in mind, the following is nevertheless a summary of our understanding of the prototypical functions of the demonstrative roots presented in Table 2–3.

- *ho?* Adverbial; extended to participants when combined with CLFs; visually or conceptually proximal (e.g. in the flow of the discourse); typically exophoric
- *m?e* Cognitively activated for speaker; assumed to be already activated for hearer; mostly endophoric and anaphoric
- *ča?a* Visually distal; typically exophoric
- *ma?a* Unseen, uncertain; inferred
- *naqae* Speaker instructs hearer to activate information that is assumed be already identifiable, known, or familiar

Relative to syntactic function, both deictic demonstratives (DDEM; with all DEM1 roots) and the recognitional demonstrative (RDEM) serve as determiners

and as participant pronominals; but only combinations with *ho?* function adverbially to signal location and time. The SDEM with *ho?* and some DDEMs with *ho?* function adverbially. It has been suggested that such a system, where the number of deictic distinctions in the pronominal domain supersedes the number of distinctions in the adverbial domain, may be comparatively rare (Levinson 2018: 19). However, the Pilagá system somewhat corresponds to the most frequent type found in Krasnoukhova's (2012) South American study, namely a system in which the same demonstrative form is used in participant-pronominal and adnominal functions (i.e. the DDEM), but not in adverbial functions (which in Pilagá mostly uses the SDEM with *ho?*). Clearly, *ho?* is a versatile element, occurring in the SDEM construction as an adverbial pro-form and in the DDEM construction for participant pronominal and determiner functions. *Naqae* functions as part of a recognitional demonstrative.

In our database, the demonstrative root tokens with exophoric function outnumber the tokens with endophoric function. Anaphoric uses are much more frequently attested than cataphoric uses. Anaphora has been pointed out as a possible source for further grammaticalisation of *m?e* as a relativiser. This development suggests that it is not always the most distal (nor proximal) member of a demonstrative system that is subject to further grammaticalisation (Himmelmann 1996: 217). What appears significant in the development of *m?e* as a relativiser is its endophoric+anaphoric profile, not a distal/proximal feature. If the subordinator *da?* is historically related to the CLF *da?* 'vertically extended', the semantic pathway must be via the extension of *da?* for abstract nominal concepts.

Corpus examination shows that essentially all determiners contain a CLF. In fact, the basic determiner is just a CLF, either deictic or postural. It would be communicatively unusual for essentially every nominal in discourse to be marked by a demonstrative; therefore we conclude that CLFs do not have the typical usage profile of demonstratives.

The extension of some CLFs and demonstrative roots into temporal and evidential/certainty meanings does not appear to be a widespread cross-linguistic feature of demonstrative systems. However, it is found in nearby Nivaclé (Gutiérrez 2015) and Wichí (Nercesian 2014: 175); in Chorote (Carol 2011); and in Movima, possibly Chapacuran Wari', and some other South American languages (Krasnoukhova 2014). The postural information found in the Guaykuruan determiner and demonstrative systems is rare, but it is also attested elsewhere, for example in the demonstrative system of the Chadic language Goemai (Hellwig 2018). The evidential/(un)certainty semantics found in the Pilagá system is connected

to speaker-anchored distance/non-visibility of referents. Evidential functions of demonstratives and determiners also appear to be typologically rare. The extent to which these relatively unusual features occur in other languages of the Chaco, South America, and beyond merits further study.

Acknowledgments

We are grateful to Pilagá speakers who have assisted with this research, especially Ignacio Silva and José Miranda. We also thank Belén Carpio, Olga Krasnoukhova, Manuel Otero, Yvonne Treis, and an anonymous reviewer for comments. This research is partially supported by NSF grant BCS 1263817 and by the Argentinian Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

Abbreviations

A	roughly active set of verbal person markers	INTG	interrogative
AUG	augmentative	INTJ	interjection
B	roughly stative set of verbal person markers	INTSF	intensifier
BCLF	basic classifier construction	IT	itive
CAUS	causative	LOC	locative
CLF	classifier	M	masculine
COL	collective	NEUT	neutral deixis
COMPL	completive	NEXIST	non-existing
CONJ	conjunction	NMLZ	nominaliser
DEM	demonstrative	NO.EXT	non-extended
DDEM	complex deictic demonstrative	NPROG	non-progressive
DES	desiderative	NVIS	non-visible
DIM	diminutive	OBJ	object
DIST	distal	PAUC	paucal
F	feminine	PL	plural
HOR	horizontal	POS	possessive
HUM	human	PROG	progressive
INDF	indefinite/nonspecific	PROX	proximal
		RCG	recognitional
		RDEM	complex recognitional demonstrative
		SBJ	subject

SDEM	simple demonstrative	VEN	ventive
SG	singular	VER	vertical
SUB	subordinator	VIS	visible

References

- Aikhenvald, Alexandra Y. 2015. Evidentials: Their links with other grammatical categories. *Linguistic Typology* 19(2). 239–277.
- Carol, Javier. 2011. Determinantes demonstrativos en chorote (mataguayo). Interrelación con la modalidad, la temporalidad y la evidencialidad. *Indiana* 28. 227–254.
- Carpio, María Belén. 2012. *Fonología y morfosintaxis de la lengua hablada por grupos tobas en el oeste de Formosa (Argentina)*. Munich: Lincom.
- Cúneo, Paola. 2016. Expresión de foco en toba (guaycurú). Una aproximación al interjuego entre sintaxis y estructura de la información. *Cuadernos de la Alfa* 9. 64–79.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 2003. Demonstratives in language use and grammar. Lecture notes of the Summer School in Functional and Cognitive Linguistics, San Marino. Unpublished ms.
- Gildea, Spike. 1993. The development of tense markers from demonstrative pronouns in Panare (Cariban). *Studies in Language* 17. 53–73.
- González, Raúl Eduardo. 2015. Demostrativos pronominales y adnominales en toba del este de Formosa (Argentina). *UniverSOS* 12. 147–162.
- Gutiérrez, Analía. 2015. Evidential determiners in Nivaclé. *Anthropological Linguistics* 57(4). 412–443.
- Hellwig, Birgit. 2018. “See this sitting one”: Demonstratives and deictic classifiers in Goemai. In Stephen C. Levinson, Sarah Cutfield, Michael J. Dunn, N. J. Enfield & Sérgio Meira (eds.), *Demonstratives in cross-linguistic perspective*, 134–149. Cambridge: Cambridge University Press.
- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Klein, Harriet. 1979. Noun classifiers in Toba. In Madeline Mathiot (ed.), *Ethnolinguistics: Boas, Sapir and Whorf revisited*, 85–95. The Hague: Mouton.
- Krasnoukhova, Olga. 2012. *The noun phrase in the languages of South America*. Utrecht: LOT.

- Krasnoukhova, Olga. 2014. The noun phrase: Focus on demonstratives, redrawing the semantic map. In Loretta O'Connor & Pieter Muysken (eds.), *The native languages of South America: Origins, development, typology*, 250–273. Cambridge: Cambridge University Press.
- Levinson, Stephen C. 2018. Introduction: Demonstratives: Patterns in diversity. In Stephen Levinson, Sarah Cutfield, Michael J. Dunn, N. J. Enfield & Sérgio Meira (eds.), *Demonstratives in cross-linguistic perspective*, 1–42. Cambridge: Cambridge University Press.
- Messineo, Cristina. 2003. *Lengua Toba (Guaycurú): Aspectos gramaticales y discursivos*. Munich: Lincom.
- Messineo, Cristina, Javier Carol & Harriet Klein. 2016. Deixis y contacto en la región del Gran Chaco: Los demostrativos en las lenguas guaycurúes y mataguayás. *International Journal of the Sociology of Language* 240. 119–157.
- Messineo, Cristina & Paola Cúneo. 2019. Multifunctionality of deictic classifiers in the Toba language (Guaycuruan). In Alexandra Y. Aikhenvald & Elena Mihas (eds.), *Classifiers and genders: A cross-linguistic perspective*, 197–221. Oxford: Oxford University Press.
- Nercesian, Verónica. 2014. *Wichi lhomtes: Estudio de la gramática y la interacción fonología-morfología-sintaxis-semántica*. Munich: Lincom.
- Nordlinger, Rachel & Louisa Sadler. 2004. Nominal tense in crosslinguistic perspective. *Language* 80. 776–806.
- Payne, Doris L. 1986. Noun classification in Yagua. In Colette Craig (ed.), *Noun classes and categorization*, 113–131. Amsterdam: Benjamins.
- Vidal, Alejandra. 1997. Noun classification in Pilagá. *Journal of Amazonian languages* 1. 58–111.
- Vidal, Alejandra. 2001. *Pilagá grammar*. Eugene, OR: University of Oregon. (Doctoral dissertation).
- Vidal, Alejandra & Analía Gutiérrez. 2010. Tiempo nominal en las lenguas chaqueñas. In Victor Castel & Liliana Severino (eds.), *La renovación de la palabra en el bicentenario de la Argentina: Los colores de la mirada lingüística*, 1347–1355. Mendoza: Universidad Nacional de Cuyo.
- Vidal, Alejandra & Harriet Klein. 1998. Irrealis in Pilagá and Toba? Syntactic versus pragmatic coding. *Anthropological Linguistics* 40(2). 175–197.

Chapter 8

Referential shift potential of demonstrative pronouns – Evidence from text continuation

Melanie Fuchs

University of Cologne

Petra B. Schumacher

University of Cologne

In this chapter, we explore different discourse functions of two types of German demonstrative pronouns compared to the personal pronoun. Utilising a continuation task, we first demonstrate that personal and demonstrative pronouns refer back to different referents from previous discourse. We then show that anaphoric pronouns can also initiate a referential shift towards a previously less prominent referent in upcoming discourse, and we compare three types of pronouns in this regard. We also demonstrate that the referential shift potential is modulated by context-dependent factors. Furthermore, we present evidence that the two demonstrative pronouns differ in the duration of their discourse structuring capacity in the unfolding text.

1 Introduction

1.1 Functions of demonstrative pronouns

One well-attested use of demonstratives is the tracking or anaphoric use, where demonstratives pick up an entity that has been introduced in previous discourse (e.g. Himmelmann 1996). In this chapter, we will investigate the anaphoric use of two types of German demonstratives. The first type encompasses the nominative singular forms *der* (M), *die* (F) and *das* (N) and the second type encompasses the



nominative singular forms *dieser* (M), *diese* (F) and *dieses* (N). There is a third type (*jener* (M), *jene* (F), *jenes* (N)) that, however, is rarely used anymore as it is considered outdated by many speakers. The demonstrative forms can be used pronominally and also occur in adnominal use (as in *der Sprachwissenschaftler*, *dieser Sprachwissenschaftler*, *jener Sprachwissenschaftler* ‘this linguist’). We are interested in the pronominal use of the more common demonstratives *der* and *dieser*.¹

Our goal for this chapter is twofold. First, we want to compare the referential preferences of demonstrative pronouns to those of personal pronouns. In other words, we want to find out what kind of referents demonstrative pronouns preferably refer back to in contrast to personal pronouns. There are many psycholinguistic studies that compare the demonstrative pronoun *der* to the personal pronoun *er*. However, there is not much empirical evidence regarding the demonstrative pronoun *dieser*.

Secondly, and more importantly, we want to investigate to what extent demonstrative pronouns influence referential chains in upcoming discourse. Based on theoretical accounts (e.g. Weinrich 1993; Abraham 2002), we hypothesise that demonstrative pronouns initiate a referential shift in upcoming discourse towards a referent that has been less prominent in previous discourse. Specifically, we are interested in the difference between the two types of German demonstratives and hypothesise that they provoke different referential dynamics in upcoming discourse. We thus assume that anaphorically used demonstrative pronouns do not only refer back to one particular (usually less prominent) entity from the previous discourse, but also promote that previously less prominent entity to a higher discourse status in upcoming discourse. There is little evidence regarding this functional component of (German) demonstrative pronouns. Therefore, we will place special emphasis on the two types of German demonstrative pronouns and their influence on story development with regard to the role of different discourse participants. In order to test our hypotheses, we conducted a story continuation task.

1.2 Choice of referent

Most research so far has tried to identify the preferred referents of personal and demonstrative pronouns. So-called accessibility or prominence hierarchies explain why personal and demonstrative pronouns refer back to different ref-

¹In this chapter we will focus on the masculine forms of these pronouns for reasons that have to do with the design of our experiment. This will be explained in detail in §2.2.

erents. According to these hierarchies, different referential forms indicate the prominence of an entity in discourse. Prominence is understood here as a relational notion that singles out one entity from a set of entities of equal type and structure (see von Heusinger & Schumacher 2019 for a comprehensive discussion of the properties of prominence). It capitalises on the competition between potential referents and is thus more refined than the static and non-relational conception of cognitive accessibility that links a referential form to a particular cognitive state (Ariel 1990; Gundel et al. 1993). While indefinite expressions refer to entities with low prominence, personal pronouns or null forms refer to entities with high prominence. Various other forms can be found in the middle part of the spectrum (Ariel 1990; 2004; Gundel et al. 1993). In German, demonstrative pronouns are placed at the higher end of the prominence scale, below unstressed personal pronouns (Ahrenholz 2007; Ellert 2011). It has been pointed out that while demonstrative pronouns refer to highly prominent (i.e. previously mentioned) entities, they explicitly avoid the most prominent entity. This observation has been made for many languages, including German, Dutch, Russian, Afrikaans and Norwegian (Johannessen 1996; Comrie 1997).

Theoretical and experimental research has tried to identify the features that contribute to a referent's prominence status in discourse, and several factors have been proposed in this context. Regarding German, the grammatical role of the referent has been discussed as an important prominence-lending factor, with the subject of a sentence being considered more prominent than the object of a sentence (Bosch et al. 2003; 2007; for Dutch, see Kaiser & Trueswell 2004). These original accounts have since been modified. For example, it has been proposed that sentence topics which often appear as the grammatical subject of a sentence are more prominent than other discourse participants (Bosch & Umbach 2007). Furthermore, the thematic agent, which initiates an action or experiences an emotion (Dowty 1991; Primus 2012), has been shown to be more prominent compared to referents with other thematic roles, such as the patient of an action (Schumacher et al. 2016). Finally, the perspectival centre and thus the person from whose perspective an event is told has been claimed to be more prominent (Hinterwimmer & Bosch 2016; Hinterwimmer 2019). The linear order of the referents itself has not significantly affected the results of any of these experiments.

It is important to note that these claims are based on the comparison between the German personal pronoun *er* and the demonstrative pronoun *der*. In most of the above cases, experimental settings were designed such that the personal or demonstrative pronoun had to be resolved towards one of two discourse referents (e.g. the thematic agent or thematic patient of a sentence, as in the study

reported by Schumacher et al. 2016). On the basis of these results, it has been concluded that the personal pronoun refers back to the most prominent subject/topic/agent whereas the demonstrative pronoun *der* refers back to the less prominent object/non-topic/patient. The ranking of these factors is still part of ongoing research.

However, there are not many investigations of the other type of demonstrative pronoun in German, namely *dieser*. A last-mentioned preference has been suggested for *dieser* (Zifonun et al. 1997), according to which *dieser* simply selects the last-mentioned candidate from the previous utterance as referent irrespective of its other features such as grammatical or thematic role. However, recent work could not confirm a last-mentioned preference for *dieser*. Rather, *dieser* seemed to pattern with *der* in preferring the patient irrespective of its sentence position (Lange 2016; Özden 2016; Patil et al. 2020). Other characteristics of *dieser* have been discussed as well. For example, it has been observed that *dieser* is used to express contrast or delimitation (Bisla-Müller 1991). This is why *dieser* is sometimes found in combination with the more antiquated *jener* (as in ‘not this one, but that one’). Some argue that *dieser* refers to the more proximal referent in such comparative constructions (e.g. Bisla-Müller 1991), but there are also diverging accounts (see Ahrenholz 2007 for an overview). All in all, there is neither a comprehensive description of *dieser* nor an account of a systematic distinction between the two types of German demonstrative pronouns *der* and *dieser* regarding their interpretive preferences. Therefore, one aim of the current study is to shed light on the resolution patterns of the two types of demonstrative pronouns.

1.3 Referential shift potential

The main goal of our study is concerned with the idea that referential expressions are used “to mark key concepts [...] that might play a pivotal role in the upcoming discourse” (Gernsbacher & Shroyer 1989: 536). According to this notion, referential expressions do not only establish links with previously mentioned entities, but also indicate to the addressee which entity will be central in upcoming discourse (see also von Heusinger & Schumacher 2019 for dynamicity as a criterion of prominence). This function has been less investigated. However, there are a few studies that illustrate the idea of referential expressions shaping the upcoming discourse. Givón (1983) edited a volume about the link between different referential expressions and their function with respect to signalling topic (dis)continuity. As part of this volume, Brown (1983) investigated a written English narrative and confirmed that the referents of shorter expressions such as zero or unstressed personal pronouns are most likely to be mentioned again in

the immediately following discourse. Gernsbacher & Shroyer (1989) specifically looked at the role of the (English) indefinite demonstrative determiner *this* in shaping upcoming discourse. They employed a story continuation task in which the participants heard the beginning of short stories. At the end of each story, a new character was introduced either with *this* or *a(n)* as determiner, as illustrated in (1).

- (1) (Gernsbacher & Shroyer 1989: 537)
- I went to the coast last weekend with Sally. We'd checked the tide schedule 'n we'd planned to arrive at low tide – 'cuz I just love beachcombin'. Right off, I found 3 whole sand dollars. So then I started lookin' for agates, but I couldn't find any. Sally was pretty busy too. She found *this/an* egg ...

The participants were then asked to continue the story. The authors report that when the indefinite demonstrative determiner *this* preceded the newly introduced character, the participants mentioned the respective character more often in their continuations. Furthermore, it was mentioned more often in the first sentence following its introduction and was referred to with less complex referential expressions. The authors therefore conclude that the demonstrative determiner *this* boosted the accessibility of the newly introduced character. As a result, the participants' story continuations developed in favour of that newly introduced character instead of the previously prominent characters. Chiriacescu (2011) reports similar results for the English indefinite *this*.

Regarding German, we assume that anaphoric demonstrative pronouns can also initiate a referential shift in upcoming discourse. This assumption is supported by several accounts in the literature describing the functions of (adnominal and pronominal) demonstratives in German. For example, it has been observed that adnominally used demonstratives function as "Aufmerksamkeits- und Warnsignale" (attention and warning signals, Weinrich 1993: 441) that indicate a change in the referential structure. Pointing in a similar direction, Abraham (2002) states that while personal pronouns continue the current discourse theme, demonstrative pronouns initiate a thematic change.

Three studies have investigated the influence of adnominally and pronominally used German demonstratives on upcoming discourse. The first (Deichsel & von Heusinger 2011) compared the German indefinite demonstrative determiner *dieser* to the indefinite determiner *ein*. Similar to the study by Gernsbacher & Shroyer (1989), the participants received short stories (in written form). A new character was introduced either with the demonstrative determiner, as in *die-*

ser Mann ‘this man’, or with the indefinite determiner, as in *ein Mann* ‘a man’. Participants were instructed to continue the story. When the new character was introduced with the demonstrative determiner, the participants more often referred to that character in the continuations. Furthermore, the participants also more frequently initiated a topic shift towards that character compared to when the new character was introduced with the indefinite determiner *ein*.

The second study (Ahrenholz 2007) focused specifically on the difference between the two types of German demonstratives *der* and *dieser* in their adnominal and pronominal use. Based on spoken corpora, the author reports that the demonstrative *der* was used to place special emphasis on a referent and to maintain that referent as the new centre of attention in upcoming discourse. The other type of demonstrative, *dieser*, was often used to single out one particular referent among many possible and similar referents. To illustrate this, the author describes a conversation about choosing one of two exam questions where pronominal *dieser* is used to specifically refer to the one that was chosen (*eine oder zwei Fragen* (F) – *eine* (F) – *diese* (F) ‘one or two questions – one (of those) – this (one)'). Finally, Schumacher et al. (2015) implemented a story continuation task in order to investigate the topic shift potential of the German demonstrative pronoun *der* compared to the personal pronoun *er*. The participants received the beginning of a story; a first sentence introduced one prominent and one less prominent character (based on their thematic roles) and a second sentence was trimmed after an ambiguous pronoun (either *der* or *er*). When the second sentence contained the demonstrative pronoun, the participants were more likely to initiate a topic change towards the previously less prominent character in their continuations.

All of the above-mentioned studies suggest that demonstratives have the potential to change the referential structure of the unfolding discourse towards previously less prominent entities (note, however, that the different studies employed different measurements to assess referential change). However, there is little evidence regarding the referential shift potential of anaphorically used demonstrative pronouns as most studies looked at demonstratives in adnominal position. Furthermore, little is known about the difference between the two types of German demonstratives and especially about the functional contribution of *dieser*. In the following study, we will therefore investigate to what extent demonstrative pronouns are understood as a signal for referential shift compared to personal pronouns. Specifically, we will investigate the difference between *der* and *dieser* with regard to their referential shift potential.

To summarise, we pursue two research goals. First, we want to determine the preferred referents of two types of German demonstrative pronouns (*der* and

dieser) when compared to the German personal pronoun. Secondly, we want to investigate to what extent German demonstrative pronouns initiate a referential shift towards previously less prominent entities in upcoming discourse. In particular, we intend to identify the forward directed functions of the German demonstrative pronouns *der* and *dieser* and potential differences between the two types of referential expressions.

1.4 The current research

To address the questions that have been outlined in the previous sections, we conducted a text continuation task (based on Gernsbacher & Shroyer 1989). Our participants received a text fragment in written form which consisted of one and a half sentences. The first sentence introduced two masculine characters. One character represented the proto-agent and one character the proto-patient of the sentence. The proto-agent of a predicate is characterised by volition, movement, causality and sentience. The proto-patient, on the other hand, is associated with undergoing a change of state and being affected by the action of the predicate (Dowty 1991; Primus 2012). The second sentence that our participants received was only a fragment and contained a masculine singular pronoun (either *er* or *der* or *dieser*) that could potentially be linked to either character from the first sentence due to its grammatical gender. Participants were then asked to continue the story by writing down six additional sentences. This way, we could analyse (i) how they understood the ambiguous pronoun in the second sentence, which is important for our research question regarding referential choice, and (ii) which character they mentioned predominantly in their continuations, which is important for our main research question regarding the referential shift potential.

1.4.1 Predictions for choice of referent

Regarding the choice of referent of the different pronouns, we have the following hypotheses. As described in §1.1, personal pronouns preferably refer to the most prominent entity, while demonstrative pronouns, crucially, are claimed to avoid the most prominent entity. It has been demonstrated in several studies that the German demonstrative pronoun *der* referred to the less prominent character (e.g. Bosch et al. 2003; 2007; Hinterwimmer & Bosch 2016). In our case, we follow previous research that has identified agentivity as a key prominence-lending cue in pronoun resolution (Schumacher et al. 2016; 2017); in fact, agen-

tivity is a driving force in other domains as well (e.g. Kretzschmar et al. 2019).² As mentioned above, Zifonun et al. (1997) proposed a last-mentioned preference for *dieser*. In our experimental setting, we only used canonical sentences where the agent precedes the patient. Therefore, we cannot test whether *dieser* prefers the last-mentioned referent because the last-mentioned referent is also the less prominent patient. We therefore simply assume that while the German personal pronoun refers back to the more prominent agent, the two demonstrative pronouns avoid the prominent character and select the other one (possibly for different reasons that we cannot disentangle in our experiment). Moreover, previous experimental research has shown that demonstrative pronouns have a strong preference to select a less prominent referent, whereas the interpretive preferences of personal pronouns are less rigid (Bosch et al. 2007; Schumacher et al. 2016; 2017). Therefore, interpretive preferences for the demonstrative pronouns are expected to be less flexible than the preference for the personal pronoun.

1.4.2 Predictions for referential shift potential

Regarding our main question of how demonstrative pronouns and personal pronouns influence the upcoming discourse with regard to referential chains, we have the following hypotheses: we expect that the German personal pronoun maintains the already established referential structure in subsequent discourse while the two demonstrative pronouns change it in such a way that the previously less prominent character becomes more central in the development of the story. Different measurements have been proposed in order to determine the central character of a story (e.g. Givón 1983; Garrod & Sanford 1988; Gernsbacher & Shroyer 1989).³ We decided to measure the discourse status of the two discourse participants in terms of how often they are mentioned in subsequent discourse (i.e. their referential persistence). According to Givón (1983: 15) “[m]ore important discourse topics appear more frequently in the register, i.e. they have a higher probability of persisting longer in the register after a relevant measuring point”. We expect that this will give us a good indication of the discourse-

²We assume that thematic roles are an important factor that influence the prominence relations in discourse. However, as one reviewer pointed out, we cannot exclude the possibility that the participants draw additional inferences and enrich the context sentences which might influence the prominence relations. This is a difficulty that all experimental studies have to face and is difficult to control for, but it crucially does not affect our main research target of identifying functional differences between the three pronominal forms.

³Among these measures are referential persistence, referential distance, immediacy of reference, referential explicitness, nature of potential competitors and topic shift potential (Givón 1983; Garrod & Sanford 1988; Gernsbacher & Shroyer 1989).

structuring potential of the different pronominal forms. In particular, the more prominent an entity is, the more probable it is that this entity is mentioned again in subsequent discourse.

As outlined above, we assume that one of the two characters from the first sentence is more prominent than the other because of its thematic features that characterise it as a proto-agent. We hypothesise that demonstrative pronouns change the prominence structure in upcoming discourse and promote the previously less prominent entity to a more prominent status in subsequent discourse. We therefore predict more references to the previously less prominent protopatient in the continuations when the participants encounter one of the demonstrative pronouns in the story fragment. In contrast, we expect the participants' story continuations to mainly centre around the prominent character when the second sentence contains the personal pronoun, since personal pronouns have been claimed to maintain the referential structure and continue the current topic (e.g. Abraham 2002).

In this context, we also test two complementary hypotheses that differ with respect to whether the referential shift potential depends on the interpretive preferences of the pronouns. As described above, the participants first had to continue the second sentence which contains an ambiguous pronoun (either *er* or *der* or *dieser*) and thus assign a referent (one of the two characters from the first context sentence) to the pronoun before they continued the story. How does the referential choice influence the referential shift potential of the different pronouns?

Hypothesis 1 views referential shift as an intrinsic property of a demonstrative pronoun (as, for instance, suggested by Weinrich 1993; Abraham 2002). In this case, the demonstrative pronouns would initiate a referential shift in subsequent discourse irrespective of whether they were interpreted as referring back to the first- or second-mentioned character from the first context sentence.

Hypothesis 2 proposes the interdependence of referential choice and the potential for referential shift (which might be suggested by Givón's (1983) approach to consider previous and upcoming discourse to determine topic continuity). This means that the referential shift potential is modulated by the referent that is chosen for the pronoun. When the demonstrative pronoun refers to the less prominent referent, this referent receives a prominence boost and is more likely to be mentioned more frequently in upcoming discourse. However, when the demonstrative pronoun refers back to the already prominent referent, there are no changes in the referential structure of the upcoming discourse, as the already prominent character does not profit from the additional prominence boost.

Finally, we are interested in the difference between the two types of German demonstrative pronouns (*der* vs. *dieser*) with regard to their referential shift potential. We expect that the two demonstrative pronouns provoke different referential dynamics in subsequent discourse. We therefore counted references to the two characters over the course of story development. For spoken German, it has been observed that *der* is often used to maintain a particular referent as the new centre of attention (Ahrenholz 2007). This might indicate that *der* is used to initiate a longer-lasting shift in the referential structure. Based on this account, we expect that the number of references to the previously less prominent character stays high over the course of story development following the demonstrative pronoun *der*. In contrast, we do not expect such a stable effect following the demonstrative pronoun *dieser*. This might be supported by Weinrich (1993: 441) who points out that *dieser* warns the addressee that there will be a “bent” (which we understand as a temporary change) in the referential structure.

2 Text continuation task

2.1 Participants

The story continuations of 112 participants (82 women, 28 men and 2 of unknown gender; mean age: 22.71 years, SD: 5.39 years) informed our analysis.⁴ They were monolingual speakers of German and mostly students from the University of Cologne who participated voluntarily or as part of coursework.

2.2 Material

For our text continuation task, we created a total of 24 incomplete pairs of context and target sentences. The first sentence introduced two masculine charac-

⁴Originally, we collected stories from 135 participants. However, we had to exclude 23 participants/stories from our analysis. Thirteen of those were excluded because they contained direct speech. We decided to exclude stories containing direct speech because the mechanisms for referential relations within direct speech might be different. Three further stories were excluded because the intended referent of the critical pronoun in the second sentence was not identifiable, one was excluded because the demonstrative pronoun *der* was understood as the definite article (despite our efforts to insert an adverb after the pronoun), four because they were ungrammatical or nonsensical, one because the animate character from the first sentence was interpreted as an inanimate object and one because the participant misunderstood the context sentence. For the analysis regarding the referential shift potential of the different pronouns, four individual data points had to be excluded because the referent of the referential expression was not identifiable. Therefore, we report the results for the 112 participants/stories that were included in our analysis.

ters (a proto-agent and a proto-patient) and the second contained one of three ambiguous masculine pronouns (either *er* or *der* or *dieser*; see (2) for examples). We decided to use masculine characters and pronouns because in German the singular feminine pronouns (*sie*, *die*, *diese*) are identical with the plural forms (*sie*, *die*, *diese*). The syncretism could have influenced the understanding of the pronouns and confounded the results, therefore masculine characters were employed in this study. The second sentence was discontinued after the ambiguous pronoun and an additional adverb. We inserted the adverb after the pronoun in order to prevent participants from understanding the demonstrative pronouns (*der*, *dieser*) as definite masculine determiners, because demonstrative pronouns and determiners are overlapping in German. The examples in (2) illustrate different incomplete sentence pairs. The context sentence (2a) is the same in all three cases, but the second sentence – see (2b)–(2d) – varies with respect to the pronoun (*er* vs. *der* vs. *dieser*) it contains. Note that we highlight the pronouns in the following sentence pairs for illustration. The pronouns were not highlighted in the questionnaires the participants received.

- (2) a. First context sentence (active accusative verb)

Jeden Morgen hat der Pfleger den Heimbewohner
 every morning has the.NOM nurse the.ACC resident
 gekämmt.
 combed

‘Every morning, the (male) nurse combed the (male) resident.’

- b. Incomplete second sentence (with personal pronoun *er*)

Dabei hat er oft ...
 during.this.process has he.PERS often
 ‘During this process, he often ...’

- c. Incomplete second sentence (with demonstrative pronoun *der*)

Dabei hat der oft ...
 during.this.process has he.DEM often
 ‘During this process, he often ...’

- d. Incomplete second sentence (with demonstrative pronoun *dieser*)

Dabei hat dieser oft ...
 during.this.process has he.DEM often
 ‘During this process, he often ...’

We also varied the verb type in the first sentence. Half of the items contained active accusative verbs ($n = 4$), which assign the first-mentioned character the

role of grammatical subject and thematic agent, and the second-mentioned character the role of object/patient in the canonical word order, as illustrated in (2). The other half of the items included so-called dative experiencer verbs ($n = 4$). Dative experiencer verbs are special in that they assign the first-mentioned character in the canonical word order the role of proto-agent/object and the second-mentioned character the role of proto-patient/subject, as illustrated in (3).

- (3) Example of an incomplete sentence pair (dative experiencer verb in context sentence)

- a. Im Hafen ist dem Segler der Urlauber aufgefallen.
at.the harbour is the.DAT sailor the.NOM tourist noticed
'At the harbour, the (male) sailor noticed the (male) tourist.'
- b. Wenig später hat er/der/dieser dann ...
shortly afterwards has he.PERS/he.DEM/he.DEM then
'Shortly afterwards, he then ...'

It has been argued that the canonical word order for dative experiencer verbs is object before subject because the first-mentioned object has the highest thematic role (Haider 1993). We included these different verb types in order to investigate whether they have an effect on the referential preferences of the pronouns. Both thematic agent and grammatical subject have been argued to be very prominent (Bosch et al. 2003; Schumacher et al. 2016). We assume that it is more difficult to interpret a pronoun in contexts with dative experiencer verbs where the thematic agent and the grammatical subject are not aligned. Alternatively, given that all context sentences have the canonical order (proto-agent before proto-patient), verb type may not have an influence (as shown in Schumacher et al. 2015; 2016), as the most prominent thematic role still appears before the less prominent role. In total, we created eight different context sentences (four for each verb type). Each context sentence was then combined with the three different pronoun conditions (*er* vs. *der* vs. *dieser*) as illustrated in (2), yielding a total of 24 incomplete sentence pairs.

2.3 Method and procedure

Each participant received only one incomplete sentence pair in written form. For example, one participant was presented with the incomplete sentence pair (2b), another participant was presented with the incomplete sentence pair (2c) and so on. Each participant was asked to continue the story by writing down six sentences. An extract of an example of a story continuation in the context of *der* is

provided in (4)–(5), translated into English. Example (4) shows the incomplete sentence pair the participant received, and (5) depicts the participant's continuation. We have added square brackets around all animate referential expressions in order to illustrate how we proceeded in annotating the data. The participants simply wrote down the sentences and were blind to the purpose of the study.

(4) Incomplete sentence pair (with demonstrative pronoun *der*)

Every morning, the male nurse combed the male resident. During this process, [he.DEM]_{Ref2} often ...

(5) Continuation by participant

... became very tranquil and [Ø]_{Ref2} started [Ø]_{Ref2} to reminisce about [his]_{Ref2} earlier life. [The nurse]_{Ref1} was sure that [the resident]_{Ref2} probably invented a lot of things, but [he.PERS]_{Ref1} didn't say anything, [Ø]_{Ref1} enjoyed witnessing that [the resident's]_{Ref2} eyes started to sparkle, that [he.PERS]_{Ref2} started [Ø]_{Ref2} to radiate from the inside out.

We annotated the story continuations using the annotation tool MMAX2 (Müller & Strube 2006). Firstly, we determined whether the participants understood the ambiguous pronoun that was presented to them as part of the incomplete second sentence as referring back to the first- or second-mentioned character from the first sentence. In most cases, the way the participants continued the second sentence containing the ambiguous pronoun allowed us to identify how they interpreted the pronoun. For instance, when the participant also mentioned the other character in the same sentence it was evident how the participant interpreted the pronoun. In some cases, the second sentence did not help us to identify the referent of the pronoun, and we had to look at the following sentences in order to find the intended referent of the ambiguous pronoun. There were a few cases ($n = 3$) where we could not track back the referent of the ambiguous pronoun at all, which we then excluded from the analysis. In (4), the index *Ref2* after the demonstrative pronoun indicates that we assume that the participant interpreted the pronoun as referring back to the second-mentioned character from the first context sentence.

After having identified the referent of the ambiguous pronoun, we marked all instances of references to animate referents in the following text. Expressions that we marked included: in/definite noun phrases, bare noun phrases, demonstrative noun phrases, proper names, personal pronouns, demonstrative pronouns, possessive pronouns, indefinite pronouns, relative pronouns, reflexive

pronouns, reciprocal pronouns and zero pronouns.⁵ In the final step, we coded which referent the expression we marked referred back to, i.e. whether an expression referred back to the first-mentioned referent from the first sentence, the second-mentioned referent from the first sentence, or another animate, newly introduced referent which we labelled as “other”. The indexes *Ref1* and *Ref2* in the above example indicate which referent from the first context sentence the referential expression referred to. When a plural expression was used, we coded exactly which referents it referred back to. We had the following options available to mark referential relations: referent 1, referent 2, referent 1 + other, referent 2 + other, referent 1 + referent 2, referent 1 + referent 2 + other, and other. For example, when a referential expression referred back to the first-mentioned character from the first sentence and another animate, newly introduced referent we selected the option “referent 1 and other”.

Every story was annotated by two independent annotators. The annotations were then compared, and disagreements were discussed. Due to this procedure of discussing every single story continuation, we did not measure inter-rater agreements. However, the two annotators mostly agreed with respect to whom a referential expression referred to.

2.4 Data analysis

In this section, we will describe how we proceeded in analysing the data. We performed two different analyses: one to find out the preferred referents of the three pronouns, and one to analyse the referential shift potential of the three pronouns.

2.4.1 Choice of referent

As mentioned above, we seek to shed light on the preferred referents of demonstrative pronouns compared to the personal pronoun. We performed a generalised linear mixed effects analysis in order to assess the relationship between pronoun type (*er* vs. *der* vs. *dieser*), preferred referent (first- vs. second-mentioned character from first context sentence), and verb type (accusative vs. dative experiencer verb), using the *lme4* package (Bates et al. 2015) in R (R Core Team 2017).

⁵In German, the use of zero pronouns is more restricted than in many other languages. We coded zero pronouns in the following cases: in coordinated sentences when the proto-agent of the verb was not explicitly mentioned (e.g. *As in the previous days, he entered his black BMW, [Ø] pulled out his mobile phone ...*), to-infinitives which are complements of verbs (e.g. *... and decided [Ø] to go out for a beer*) and non-finite clauses (e.g. *Later he then went to the sailor in order [Ø] to get to know him*).

We compared two models in order to investigate the effect of verb type on the choice of referent. The outcome variable in both cases was reference to the first- or second-mentioned referent. In the first model, we specified pronoun type as a fixed effect. In the second model, we added the interaction of pronoun type and verb type. Random intercepts were estimated for different items in both models. We then compared the two models using a likelihood ratio test (which was performed in R with the `anova()` function).⁶

2.4.2 Referential shift potential

As outlined in §1.4.2, we assume that we have evidence of a referential shift when the participants mention the previously less prominent character (i.e. *Ref2* from the context sentence) more often in their continuations. As an example, see (5), where all animate referents in the continuation are marked with square brackets. We used the dichotomous measure of whether each animate referent in the story referred back to the second-mentioned character (or not) as the dependent variable in our models.⁷ This way, we could estimate the likelihood that any of these referents referred to the second-mentioned character.

As we pointed out, we are specifically interested in how the different pronouns (*er* vs. *der* vs. *dieser*) influence the story continuations. We further wanted to account for the fact that the participants had to assign a referent to the pronoun before they proceeded with their story continuation as illustrated in (4). In previous studies (e.g. Gernsbacher & Shroyer 1989), by contrast, the linguistic markers were already disambiguated, as in *this egg* in (1). We therefore firstly had to determine, based on the participants' story continuation, who the pronoun referred to. In (5), the nature of the continuation makes it clear that the pronoun *der* refers to the male resident (the second-mentioned referent from the context sentence). This enabled us to answer the following two questions in our analysis: given that the participants understood the pronoun (*er*, *der* or *dieser*) as

⁶The following models were compared:

```
glmer(criticalPronounAntecedent ~ criticalPronoun + (1|itemId), data = df_backward, family = binomial(link = "logit"))
```

```
glmer(criticalPronounAntecedent ~ criticalPronoun * verbType + (1|itemId), data = df_backward, family = binomial(link = "logit"))
```

⁷Plural expressions referring to two or three referents were counted as two or three mentions. For example, when a plural pronoun referred to the first- and second-mentioned character from the first sentence, it entered our analysis as one mention of the first-mentioned character and one mention of the second-mentioned character. The number of mentions is therefore higher than the number of referential expressions that were used in the continuations by the participants.

referring to the second-mentioned (i.e. less prominent) referent, how likely is it that any animate referent in the continuations refers to the second-mentioned character? And, given that the participants understood the pronoun as referring to the first-mentioned (i.e. more prominent) referent, how likely is it that any animate referent in the continuations refers to the second-mentioned referent?

We performed a generalised linear mixed model analysis. The fixed effects were pronoun type (*er*, *der* or *dieser*) and referent of the pronoun (*Ref1* or *Ref2*). Similar to the previous analysis, we compared two models in order to investigate whether verb type has an effect on the referential shift potential of the different pronouns. In the first model, we specified pronoun type (*er* vs. *der* vs. *dieser*) and preferred referent (of the pronoun in the second sentence: *Ref1* vs. *Ref2*) as fixed effects and also included an interaction effect between the two. In the second model, we added a three-way interaction of pronoun type, preferred referent and verb type (*accusative* vs. *dative experiencer verb*). We then compared the two generalised linear mixed models using a likelihood ratio test (which was performed in R with the `anova()` function). As random effects, we specified random intercepts for items and participants in both models.⁸

2.5 Results

2.5.1 Choice of referent

Figure 1 shows the referential preferences for the three different types of pronouns. It demonstrates that the personal pronoun prefers the first-mentioned referent (in 65% of all cases) while the two demonstrative pronouns prefer the second-mentioned referent (in 73% of all cases for *der* and 74% for *dieser*). A comparison between the model including pronoun as fixed effect and the one including the interaction of pronoun and verb type did not show a significant difference in model fit (likelihood-ratio test: $p = 0.71$). This indicates that verb type did not have an influence on the results. In the following we therefore report the results for the reduced model with pronoun as fixed effect only. The model suggests that the choice of the referent significantly depends on the pronoun type. In particular it revealed a difference between *er* and *der* (difference measured in logits: -2.64, SE = 0.72, $p < 0.001$) and between *er* and *dieser* (dif-

⁸The following models were compared:

```
glmer(antecedent_ref2 ~ criticalPronounAntecedent * criticalPronoun + (1|itemId) + (1|participantId), data = df_all, family = binomial(link = "logit"))
glmer(antecedent_ref2 ~ criticalPronounAntecedent * criticalPronoun * verbType + (1|itemId) + (1|participantId), data = df_all, family = binomial(link = "logit"))
```

ference measured in logits: -2.51, SE = 0.72, $p < 0.001$). As indicated by Figure 1, the difference between *der* and *dieser* was not significant (difference measured in logits: 0.13, SE = 0.67, $p = 0.85$).

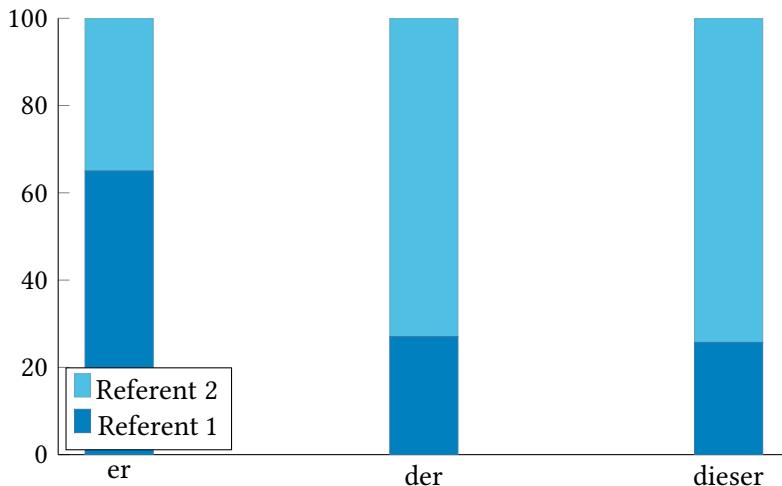


Figure 1: Percentage of reference resolution to the first- and second-mentioned referent from the first sentence for each pronoun type

2.5.2 Referential shift potential – Remention capacity

Figure 2 shows how likely it is that any mention of an animate referent in the continuations refers back to the less prominent (i.e. second-mentioned) referent from the first context sentence. The figure shows the results split by pronoun type (*er* vs. *der* vs. *dieser*). More specifically, the left panel shows the results obtained in cases when the pronoun was interpreted towards the first-mentioned character from the first context sentence; the right panel shows the results obtained in cases when the pronoun was interpreted towards the second-mentioned character in the first context sentence. When the pronouns were interpreted as referring to the less prominent character (right panel), all pronouns initiated a referential shift in subsequent discourse towards the previously less prominent character. These results hold even for the personal pronoun that has been claimed to maintain the current referential structure (e.g. Abraham 2002). However, when the pronouns were interpreted as referring to the first-mentioned character, (quantitative) differences between the different pronouns become apparent. When the demonstrative pronoun *der* was interpreted as referring to the first-mentioned

character, the second-mentioned character nevertheless appears to be accessible to a certain degree in the story continuations. By contrast, when the demonstrative pronoun *dieser* was interpreted as referring to the first-mentioned character, the second-mentioned one appears to be less accessible in the story continuations. This (quantitative trend) difference between *der* and *dieser* suggests that *der* has a slightly higher potential to change the prominence relations in discourse even when it is interpreted with respect to the more prominent referent. Similar to the analysis regarding the preferred referent of the pronouns, verb type did not significantly change the model results and is thus not represented in Figure 2.

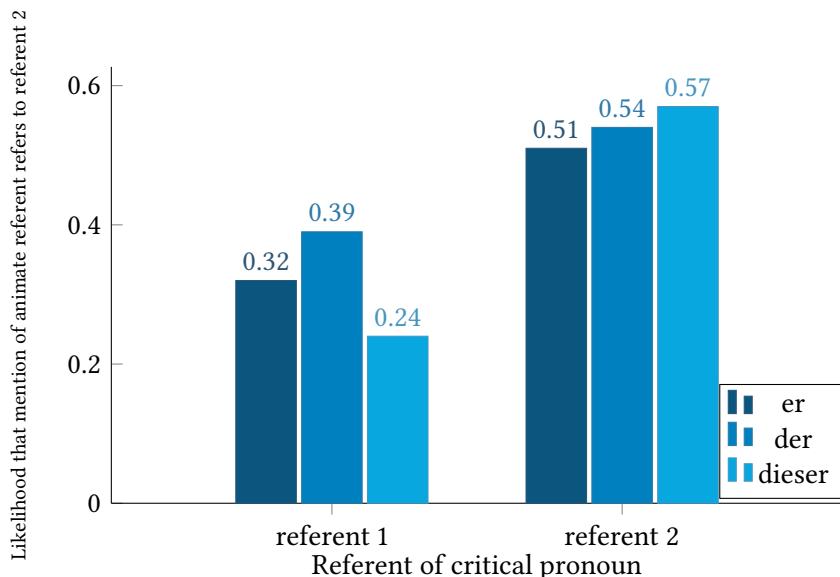


Figure 2: Probabilities (out of 1) that any mention of an animate referent in story continuations refers to the second-mentioned referent from the first sentence (depending on pronoun type and referent of pronoun)

The comparison between the model with a pronoun type by preferred referent interaction and the one containing the three-way interaction of preferred referent, pronoun type and verb type did not reveal a significant improvement of the model fit (likelihood-ratio test: $p = 0.17$).⁹

⁹R returned the warning message that the more complex model with the three-way interaction did not converge. The results of the model comparison should therefore be interpreted with caution. To double check, we ran the two models as Bayesian mixed models (using the R

Statistical analyses of the predicted probabilities are thus reported for the model containing the interaction of pronoun type and preferred referent as fixed effects and random intercepts for items and participants. Participants were significantly more likely to choose to mention the second-mentioned character if they interpreted the critical pronoun as referring to the second-mentioned character from the first context sentence. This is indicated by a model with the same dependent variable (i.e. whether an animate referent in the continuations refers to *Ref2* or not) and the variable referent of the critical pronoun (i.e. who they assumed the ambiguous pronoun referred to) as the only fixed effect (difference measured in logits: 0.93, SE = 0.19, $p < 0.001$). The differences within the two panels that are visible in Figure 2 did not reach statistical significance (possibly due to insufficient sample size).

2.5.3 Referential shift potential – Referential dynamics

As specified in §1.4.2, we were also interested in whether the two demonstrative pronouns evoke different referential dynamics. Figure 3 illustrates how often the first- or second-mentioned referent is mentioned (in absolute numbers) at different points in the stories depending on the pronoun type.

The top part of the figure illustrates how often the first-mentioned character is mentioned throughout the stories depending on the pronoun type. The x-axis refers to different points in the story. More specifically, it refers to the n^{th} mention of any animate referent in the story continuations. We enumerated all references to animate entities in the stories. The two animate characters from the first context sentence are the first two mentions and the pronoun in the second sentence is the third mention. The figure begins with the fourth mention of an animate entity and thus the first time the participants mentioned an animate referent. In the story continuation from (5), this would be the zero-pronoun marked on *started*. For example, in about five cases across all stories where the target sentence contained *dieser*, the fourth mention of an animate character refers to the first-mentioned character. For all pronouns, the number of references to the two characters declines towards the end. In this context, it is important to note that the story continuations were of different length and therefore contained a different number of mentions of animate referents.

Overall, the top half of the figure suggests that the personal pronoun *er* (dashed line) evokes a higher number of references to the first-mentioned character over

package brms (Bürkner 2017) with default priors) and compared them using Bayes factors. The results of this comparison also provide evidence for the reduced model without the three-way interaction ($\text{BF} = 0.002$).

the course of the stories compared to the two demonstrative pronouns. This is especially clear from the beginning up until the 12th mention of an animate referent.

The lower part of the figure demonstrates how often the second-mentioned character from the context sentence is mentioned. Compared to the top part of the figure, it suggests that both types of demonstrative pronouns activate the second-mentioned character from the context sentence more often than the first-mentioned entity. Furthermore, the two demonstrative pronouns appear to boost reference to the second-mentioned character at different points in the stories. The number of mentions of the second-mentioned character in the context of *dieser* (solid line) peaks at an earlier point in the development of the story and then declines. By contrast, in the context of *der* (dotted line) the number of mentions of the second-mentioned character has an initial peak around the sixth mention and reaches its climax at a later point and, importantly, remains relatively high in the following period.

3 General discussion

Our goal for this paper was two-fold. First, we wanted to compare the referential preferences of demonstrative pronouns to those of personal pronouns (choice of referent). Secondly, we wanted to investigate how far demonstrative pronouns initiate a referential shift in the following discourse (referential shift potential).

3.1 Choice of referent

The results for the interpretive preferences of the three types of pronouns investigated are fairly straightforward (see Figure 1). The personal pronoun shows a preference for the first-mentioned proto-agent argument, and both demonstrative pronouns reject this referent for the most part and have a robust preference for the second-mentioned proto-patient argument. These findings are in line with previous research on *er* and *der* (e.g. Bosch et al. 2003; Schumacher et al. 2016). The results further provide new data on the interpretive preference of *dieser*. Crucially, in context sentences with two referents and canonical argument order (proto-agent before proto-patient), the two demonstrative pronouns pattern alike. Whether the resolution preferences for *dieser* are triggered by an anti-agent preference (as, for example, shown by Schumacher et al. 2016 for *der*) or a last-mentioned preference (Zifonun et al. 1997) cannot be determined on the basis of our experimental design. Note, however, that recent investigations

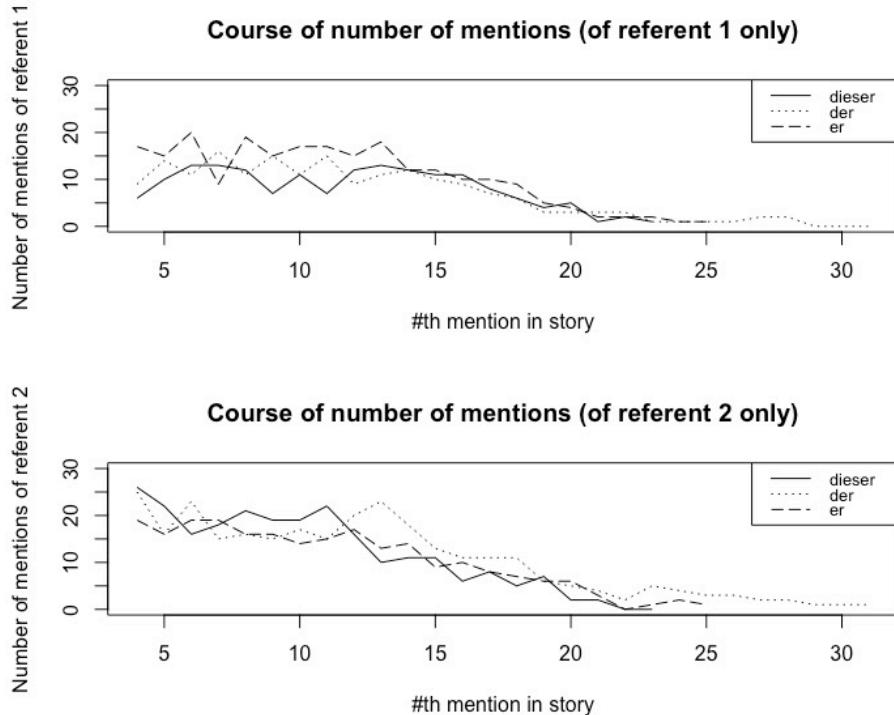


Figure 3: Number of times (absolute numbers) that the first- or second-mentioned character from the first sentence is mentioned at the n^{th} mention of any animate referent in story continuations

question the last-mentioned preference that has been suggested for *dieser* (Lange 2016; Özden 2016; Patil et al. 2020). In these experiments, contexts with non-canonical argument order were tested, as illustrated in (6) (taken from Özden 2016).

- (6) Den Direktor will der Schüler begrüßen.
 the.ACC principle wants the.NOM student to.welcome
 ‘The student wants to welcome the director.’

In German non-canonical sentences, the less prominent thematic role (protopatient) appears before the agent. According to the last-mentioned account, *dieser* would simply select the last-mentioned entity which in this case is the agent. However, the participants in all three studies interpreted both demonstrative

pronouns as referring to the first-mentioned patient. These results contradict a last-mentioned account and favour an account that is centred around thematic roles.

3.2 Referential shift potential

With regard to our main research question of how pronouns influence the upcoming discourse, we were interested in two different aspects: we analysed how likely it is that any mention of an animate referent refers to the previously less prominent character, and we looked at the referential dynamics of the unfolding stories.

The results regarding how often the two discourse participants from the first context sentence were picked up in the continuations indicate differential effects. Which character is more likely to be mentioned in subsequent discourse appears to be influenced by the referent that was chosen for the pronoun (see Figure 2). When the participants interpreted the pronoun as referring back to the less prominent character (see Figure 2, right panel), this character appears to be more likely to be mentioned more frequently in the subsequent discourse – irrespective of which pronoun the second context sentence contained. This suggests that when a less prominent character is chosen as referent for a subject pronoun and thus a referential shift is initiated, this has consequences for the development of subsequent discourse. This was expected for the two types of demonstrative pronouns, but not for the personal pronoun. However, personal pronouns have been shown to be generally more flexible in their referential choice (e.g. Bosch et al. 2007; Schumacher et al. 2016). Therefore, it might not be surprising that they can initiate a referential shift in the few cases when they were initially understood as referring back to the less prominent character (remember that in around 35% of the cases the personal pronoun from the second context sentence was understood as referring back to the less prominent patient, as illustrated in Figure 1, and that only in these cases did the personal pronoun initiate a referential shift in upcoming discourse). This suggests that the commitment to the less prominent referent was stronger overall.

However, when the participants interpreted the pronouns as referring back to the more prominent character, quantitative differences between the different pronoun conditions became apparent (see Figure 2, left panel). Interestingly, the demonstrative pronoun *der* appears to have a small potential to shift the referential focus to the less prominent patient even when it was initially interpreted towards the more prominent agent. In contrast, the demonstrative pronoun *dieser* does not show such an effect. We have to point out that these differences

within the left panel of Figure 2 are statistically not significant, but display an interesting numerical trend.

These findings provide initial support for an interdependence between interpretive preference and discourse change potential. The referential shift potential of different pronouns appears to be modulated by their interpretive preference in the first instance. It is thus not true across the board that demonstrative pronouns initiate a referential shift¹⁰ and that personal pronouns maintain the previously established prominence ranking. When the personal pronoun is interpreted as referring to the less prominent character, it can initiate a referential shift in favour of this less prominent character. Similarly, when the demonstrative pronoun *dieser* is interpreted as referring to the more prominent character, it does not necessarily initiate a referential shift but might continue with the more prominent character in subsequent discourse. Based on our results, we therefore argue for a more differentiated view when it comes to describing the referential shift potential of different pronouns or other linguistic markers.

Furthermore, the close investigation of referential dynamics (Figure 3) across the story revealed interesting results. Different referential dynamics for the three pronouns could be observed. In the context of the demonstrative pronoun *dieser*, the number of references to the less prominent character increased only for a short period of time. This mirrors the characterisation of *dieser* as introducing a momentary interruption of a referential chain (Weinrich 1993). In the context of *der*, the number of references to the less prominent character remained relatively high throughout the story. This corroborates findings from spoken German where *der* was used to establish a new centre of attention (Ahrenholz 2007).

3.3 Strengths and limitations

While the referential shift potential can also be investigated through corpus research (see, for example, Prince's (1981) corpus analysis of indefinite *this*), the written or spoken version of the text continuation task allows for a more controlled approach to the study of different types of referential expressions. Given its experimental setup, the context sentences can be constructed in minimal pairs and other discourse-based factors can be kept stable. Moreover, since the use of certain referential forms in natural contexts may be too infrequent to draw conclusions about their functional contributions, a controlled experimental design allows the generation of sufficient data. In addition, in contrast to other more

¹⁰See also Weinert (2011: 73), who observed that demonstrative pronouns "are not rhematic *per se*".

controlled approaches like reading time measures or eye tracking, this particular experimental task can be carried out without any additional equipment (in the written version) or with the aid of audio recording equipment (in the spoken version). It may thus serve as a valuable tool to test hypotheses linked to the discourse structuring potential of demonstratives and other discourse markers in settings where little technical equipment is available or for languages where large corpora are not available.

Yet such a controlled setup is also subject to certain caveats. First, the contextual settings in the present experimental design introduced only two discourse referents (as has been the case in most experimental studies on pronoun resolution). Both demonstrative pronouns patterned alike and were interpreted as referring back to the less prominent proto-patient. However, future research should introduce contexts with more than two discourse referents in order to reveal potential differences between the two demonstrative pronouns that we could not capture with our reduced experimental setting (like reference to a less prominent referent vs. to the last-mentioned referent). A context with more than two discourse referents might also be informative with regard to our second research question of how the two demonstrative pronouns influence referential chains in subsequent discourse. As our context sentences make available only two referential candidates, the continuation task encourages participants to tell a story about these two individuals. As a result, likelihood of remention is high for both of these referents and participants often switch back and forth between the dominant characters in the story. Future research should therefore examine contexts with a larger set of potential candidates in order to see how the pronouns influence referential chains when more than two discourse referents are accessible.

Another important issue is the question of whether the two types of demonstratives occur primarily in opposing registers and modalities and, following from this, how far our results can be applied to other contexts. In contrast to native speakers' intuitions – that the demonstrative pronoun *dieser* mainly occurs in written and/or formal contexts and *der* is used in spoken and/or colloquial contexts – both types of demonstrative pronouns are reported to surface in spoken interactions between lecturers and students (Ahrenholz 2007; Weinert 2007). These results show that *dieser* occurs in spoken contexts and that *der* is not limited to informal contexts, as interactions between lecturers and students might be considered rather formal. Weinrich (1993) also argues that demonstratives from the *der* paradigm should not be considered informal or colloquial. He points out that certain text types make the occurrence of certain types of referential expressions more likely (see also Ahrenholz 2007 for a similar account). By contrast,

in a psycholinguistic experiment where participants had to insert a personal or demonstrative pronoun into a formal or informal text segment, Patil et al. (2020) found that participants preferred *dieser* over *der* in formal written contexts. Thus, there is so far not enough evidence regarding the role that different registers and modalities play in pronoun use.

Hence the present findings may only be limited to the use of the two demonstratives in relatively formal written contexts. The two demonstratives may thus behave slightly differently in discrete contexts and modalities, but the available studies do not suggest that a preference for different registers and modalities alone can account for the differences between *der* and *dieser*. Therefore, follow-up studies are required. As a first step, a similar story continuation task in oral modality should be conducted where the participants hear the initial text segment and are then asked to continue the story orally. While experiments like the one presented in this chapter allow for a controlled approach to studying language comprehension and use, it would additionally be desirable to substantiate the findings on the basis of more natural, less controlled interactions between two or more discourse participants in written (e.g. chat rooms) and spoken contexts.

4 Conclusion

We have presented evidence from a story continuation task for different discourse functions of two types of German demonstrative pronouns compared with the personal pronoun. Regarding their interpretive preferences, the two types of demonstrative pronouns did not differ in their choice of referent and both selected the less prominent entity (i.e. the proto-patient). Our data further support previous findings that identified thematic role information to be a stronger cue during pronoun resolution than grammatical function (Stevenson et al. 1994; Schumacher et al. 2016, 2017), as in contexts with dative experiencer verbs the demonstrative pronouns referred back to the proto-patient/subject (and not the proto-agent/object).

With respect to their referential shift potential, the two demonstrative pronouns revealed distinct patterns. The demonstrative pronoun *der* showed a more robust referential shift potential. This became apparent from two observations. First, the demonstrative pronoun *der* appears to be slightly more likely to initiate a shift towards the previously less prominent character even when it was initially understood as referring back to the more prominent character. Secondly, it showed a higher number of references to the less prominent character through-

out the story. *Dieser*, by contrast, appears to have a short-lived referential reorienting capacity. Further, we provided evidence that the referential shift potential is not an intrinsic property of demonstrative pronouns but is modulated by context-dependent factors such as to whom the demonstrative pronoun refers. Importantly, referential shift can also be displayed by the personal pronoun in those cases when it was initially interpreted as referring back to the less prominent character (though in these cases the participants might have processed it as a stressed personal pronoun).

Acknowledgments

This research has been funded by the German Research Foundation (DFG) as part of the SFB 1252 “Prominence in Language” (project number 281511265) in the project C07 “Forward and Backward Functions of Discourse Anaphora” at the University of Cologne. We would like to thank Julia Plechatsch, Hilde Gschossmann, Julia Veth and Celine Cuma for help with data collection and annotation. Furthermore, we would like to thank Eric Engel for setting up our annotation software and helpful suggestions concerning data analysis as well as Fahime Same for support with data extraction and Maximilian Hörl for support with data analysis. Furthermore, we would like to thank two anonymous reviewers and the editors of the volume for helpful comments and suggestions.

References

- Abraham, Werner. 2002. Pronomina im Diskurs: deutsche Personal- und Demonstrativpronomina unter „Zentrierungsperspektive“. Grammatische Überlegungen zu einer Teiltheorie der Textkohärenz. *Sprachwissenschaft* 27(4). 447–491.
- Ahrenholz, Bernt. 2007. *Verweise mit Demonstrativa im gesprochenen Deutsch: Grammatik, Zweitspracherwerb und Deutsch als Fremdsprache*. Berlin: De Gruyter.
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Ariel, Mira. 2004. Accessibility marking: Discourse functions, discourse profiles, and processing cues. *Discourse Processes* 37(2). 91–116.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1). 1–48.
- Bisle-Müller, Hansjörg. 1991. *Artikelwörter im Deutschen: Semantische und pragmatische Aspekte ihrer Verwendung*. Tübingen: Niemeyer.

- Bosch, Peter, Graham Katz & Carla Umbach. 2007. The non-subject bias of German demonstrative pronouns. In Monika Schwarz-Friesel, Manfred Consten & Mareile Knees (eds.), *Anaphors in text: Cognitive, formal and applied approaches to anaphoric reference*, 145–164. Amsterdam: Benjamins.
- Bosch, Peter, Tom Rozario & Yufan Zhao. 2003. Demonstrative pronouns and personal pronouns: German *der* vs *er*. In *Proceedings of the 2003 EACL Workshop on the Computational Treatment of Anaphora*, 61–68.
- Bosch, Peter & Carla Umbach. 2007. Reference determination for demonstrative pronouns. In Dagmar Bittner & Natalia Gagarina (eds.), *Proceedings of the Conference on Intersentential Prenominal Reference in Child and Adult Language, December 1–2, 2006, Berlin*, vol. 48, 39–51. Berlin: Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS).
- Brown, Cheryl. 1983. Topic continuity in written English narrative. In Talmy Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 313–341. Amsterdam: Benjamins.
- Bürkner, Paul-Christian. 2017. **brms** : An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). 1–28.
- Chiriacescu, Sofiana Iulia. 2011. Effects of reference form on frequency of mention and rate of pronominalization. In Iris Hendrickx, Sobha Lalitha Devi, António Horta Branco & Ruslan Mitkov (eds.), *Anaphora processing and applications: 8th Discourse Anaphora and Anaphor Resolution Colloquium, DAARC 2011, Faro, Portugal, October 6–7, 2011. Revised selected papers* (Lecture Notes in Computer Science 7099), 132–143. Berlin: Springer.
- Comrie, Bernard. 1997. Pragmatic binding: Demonstratives as anaphors in Dutch. In Matthew L. Juge & Jeri L. Moxley (eds.), *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society, February 14–17, 1997. General session and parasession on pragmatics and grammatical structure*, 50–61. Berkeley, CA: Berkeley Linguistics Society.
- Deichsel, Annika & Klaus von Heusinger. 2011. The cataphoric potential of indefinites in German. In Iris Hendrickx, Sobha Lalitha Devi, António Horta Branco & Ruslan Mitkov (eds.), *Anaphora processing and applications: 8th Discourse Anaphora and Anaphor Resolution Colloquium, DAARC 2011, Faro, Portugal, October 6–7, 2011. Revised selected papers* (Lecture Notes in Computer Science 7099), 144–156. Berlin: Springer.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.
- Ellert, Miriam. 2011. *Ambiguous pronoun resolution in L1 and L2 German and Dutch*. Nijmegen: Radboud University. (PhD dissertation).

- Garrod, Simon & Tony Sanford. 1988. Thematic subjecthood and cognitive constraints on discourse structure. *Journal of Pragmatics* 12(5–6). 519–534.
- Gernsbacher, Morton Ann & Suzanne Shroyer. 1989. The cataphoric use of the indefinite *this* in spoken narratives. *Memory & Cognition* 17(5). 536–540.
- Givón, Talmy (ed.). 1983. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: Benjamins.
- Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307.
- Haider, Hubert. 1993. *Deutsche Syntax, generativ: Vorstudien zur Theorie einer projektiven Grammatik*. Tübingen: Narr.
- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Hinterwimmer, Stefan. 2019. Prominent protagonists. *Journal of Pragmatics* 154. 79–91.
- Hinterwimmer, Stefan & Peter Bosch. 2016. Demonstrative pronouns and perspective. In Patrick Georg Grosz & Pritty Patel-Grosz (eds.), *The impact of pronominal form on interpretation*, 189–220. Berlin: De Gruyter.
- Johannessen, Janne Bondi. 1996. DENNE. *Norsk Lingvistisk Tidsskrift* 14(1). 3–27.
- Kaiser, Elsi & John C. Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94(2). 113–147.
- Kretzschmar, Franziska, Tim Graf, Markus Philipp & Beatrice Primus. 2019. An experimental investigation of agent prototypicality and agent prominence in German. In Anja Gattnar, Robin Hörnig, Melanie Störzer & Sam Featherston (eds.), *Proceedings of Linguistic Evidence 2018: Experimental data drives linguistic theory*, 101–123. Tübingen: University of Tübingen.
- Lange, Annalena. 2016. *Interpretation von Demonstrativpronomen im Kontext von Dativ-Experiencer-Konstruktionen*. Cologne: University of Cologne. (B.A. thesis).
- Müller, Christoph & Michael Strube. 2006. Multi-level annotation of linguistic data with MMAX2. In Sabine Braun, Kurt Kohn & Joybrato Mukherjee (eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods*, 197–214. Frankfurt (Main): Lang.
- Özden, Filiz. 2016. *Interpretationspräferenzen von dieser: Eine empirische Studie*. Cologne: University of Cologne. (B.A. thesis).
- Patil, Umesh, Peter Bosch & Stefan Hinterwimmer. 2020. Constraints on German *diese* demonstratives: Language formality and subject-avoidance. *Glossa* 5(1). 14.

- Primus, Beatrice. 2012. *Semantische Rollen*. Heidelberg: Winter.
- Prince, Ellen F. 1981. On the inferencing of indefinite-*this* NPs. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (eds.), *Elements of discourse understanding*, 231–250. Cambridge: Cambridge University Press.
- R Core Team. 2017. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org>.
- Schumacher, Petra B., Jana Backhaus & Manuel Dangl. 2015. Backward- and forward-looking potential of anaphors. *Frontiers in Psychology* 6. 1746.
- Schumacher, Petra B., Manuel Dangl & Elyesa Uzun. 2016. Thematic role as prominence cue during pronoun resolution in German. In Anke Holler & Katja Suckow (eds.), *Empirical perspectives on anaphora resolution*, 213–240. Berlin: De Gruyter.
- Schumacher, Petra B., Leah Roberts & Juhani Järvikivi. 2017. Agentivity drives real-time pronoun resolution: Evidence from German *er* and *der*. *Lingua* 185. 25–41.
- Stevenson, Rosemary J., Rosalind A. Crawley & David Kleinman. 1994. Thematic roles, focus and the representation of events. *Language and Cognitive Processes* 9(4). 519–548.
- von Heusinger, Klaus & Petra B. Schumacher. 2019. Discourse prominence: Definition and application. *Journal of Pragmatics* 154. 117–127.
- Weinert, Regina. 2007. Demonstrative and personal pronouns in formal and informal conversations. In Regina Weinert (ed.), *Spoken language pragmatics: An analysis of form-function relations*, 1–28. London: Continuum.
- Weinert, Regina. 2011. Demonstrative vs. personal and zero pronouns in spoken German. *German as a Foreign Language* 1. 71–98.
- Weinrich, Harald. 1993. *Textgrammatik der deutschen Sprache*. Mannheim: Dudenverlag.
- Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Berlin: De Gruyter.

Chapter 9

Psychologically distal demonstratives in Scandinavian are not “discourse new”

Janne Bondi Johannessen

University of Oslo

Scandinavian languages have a demonstrative (in the form of the third person pronoun) used with nouns or noun phrases, which is different from more well-known spatial demonstratives. Johannessen (2008) argued that its conditions of use are related to psychological distance, while Lie (2010) argued that its main role is to invoke a referent in discourse. In this chapter, I have gone through empirical data, two short stories and four speech corpus dialogues, and investigated the use of this demonstrative (called the psychologically distal demonstrative or PDD). It is concluded that there are many occurrences of the PDD that would remain unexplained in Lie’s account: It can occur more than once per discourse, it can be used by both interlocutors in the same discourse, and not all referents are denoted by it. Also, it does not point out only key referents. An account based on psychological distance can explain the empirical facts.

1 Introduction

Johannessen (2008) describes a type of demonstratives found in the Scandinavian languages, which are called “psychologically distal demonstratives” (PDDs). They are mostly found in informal speech, and are not used in written text unless they represent a spoken discourse.

Formally, the PDD has the form of the 3rd person singular personal pronoun (*han* ‘he’ and *hun* ‘she’), used attributively to (modified) nouns. Plural usage has not been discussed in the literature. Its function, according to Johannessen (2008), is to signal psychological distance: that the speaker does not personally know the person referred to; or that the addressee does not know the person referred to; or that the speaker has a negative attitude to the person referred to. Lie (2010)



contests Johannessen's account, and instead proposes that these demonstratives are used for background deixis, referring to what Diessel (1999) calls "discourse new and hearer old". A similar understanding is found in Teleman et al. (1999: vol. 2: 317) for Swedish: to actualise referents who are not present in a concrete discourse, but who both the speaker and hearer have in their world of concepts. An example of the PDD is given in (1).

- (1) Er like gammel som **hu dama** og kunne ALDRI klart å
am equally old like she woman.DEF and could never managed to
være sammen med en så gammel mann.
be together with a such old man
'I'm as old as that woman and could never have coped being together
with a man that old.'¹

Example (1) is from a comment on a discussion forum, discussing the age difference between an older celebrity cook and his young female partner. The commentator clearly distances herself from the woman.

In this chapter, I will argue against Lie's point. Using empirical data, specifically dialogues in fictional literature as well as authentic dialogues from speech corpora, I will show that psychologically distal demonstratives are often used repeatedly throughout a dialogue. Also, I will show that the PDD can be used for the same referent by both interlocutors in a dialogue. Both uses would be hard to explain using Lie's account, while they fit well with Johannessen (2008). Although the present chapter uses Norwegian data, Johannessen (2008) deals with more Scandinavian languages, and it will be assumed, based on my knowledge of these, that what is said about Norwegian here carries over to the other languages.

The chapter is structured as follows: §2 presents the two different analyses of the PDD and a brief review of relevant accounts in the literature. §3 is a presentation of the empirical material and the method used in this study, while §4 and §5 go through findings from two short stories and two corpora used. In §6, the results are summarised and the chapter is concluded.

¹<https://forum.kvinneguiden.no/topic/1275529-hellstr%C3%B8m-og-anita/>.

2 Two different views on the Scandinavian PDD and related discussions

2.1 Psychologically distal demonstratives (Johannessen 2008)

Johannessen (2008) shows that in addition to spatial deixis, the Scandinavian languages (Norwegian, Swedish and Danish, as well as Icelandic) also have demonstratives, with the same form as 3rd person singular pronouns, that express emotional deixis, not previously described in depth. Although they have the form of pronouns, their use is different not only in that they always modify a noun and never replace a noun phrase, but also in that there are special pragmatic conditions for their use (see below). The PDD is also different from the preproprietary article found in many dialects, as explained in Johannessen (2008: 169–170):² It is generally not inflected for case; it carries stress; it has a distinct form in many dialects; it adds additional meaning to the noun phrase; it can modify any kind of noun, not just proper nouns; and is not obligatory.

Since the demonstrative co-occurs with a noun phrase, it is possible to ask whether what we see is rather a pronoun followed by an appositive noun phrase. The answer is clearly no. If the noun phrase following the demonstrative were an apposition, one would expect it to be a full noun phrase. In Norwegian, it is easy to see whether this is the case, since definite noun phrases modified by an adjective or a relative clause have obligatory double definiteness, meaning that the noun must be marked by a definite suffix and that an additional definite determiner must occur phrase-initially. This would be the noun phrase used with appositions, which is not what we see with the PDD. Instead, the PDD is an integrated part of the noun phrase (see Johannessen 2008: 185–187). A pronoun with an apposition is illustrated in (2), while the PDD is given in (3). Note also that the pronoun is inflected for case, unlike the demonstrative.³

- (2) Jeg beundrer **henne**, den flinkeste jenta i klassen.
 I admire her.ACC the cleverest.DEF girl.DEF in class.DEF
 ‘I admire her, the cleverest girl in the class.’

²A preproprietary article is a prenominal definite article that must appear with proper names in argument positions in many Scandinavian dialects (Johannessen & Garbacz 2014).

³There is, somewhat surprisingly, one single case-inflected demonstrative in the material, see (19). I will not discuss this further here, but simply note its existence.

- (3) Jeg beundrer **hun** flinkeste jenta i klassen.
I admire she.NOM cleverest.DEF girl.DEF in class.DEF
'I admire that cleverest girl in the class.'

With spoken language dialogue corpora as empirical evidence, I argue that these demonstratives regulate the degree of psychological distance the speaker wants to convey to the listener about a referent. The conditions for use below and the examples (4)–(9) illustrating such psychologically distal demonstratives are mainly taken from Johannessen (2008: 164–167). For information on informant codes, see §3.

PDD Condition 1: The speaker does not personally know the person referred to.

- (4) (NoTa_027)
Hun dama hun blei jo helt nerd da.
she woman.DEF she became yes totally nerd then
'But that woman became a total nerd, as you know.'⁴

- (5) (NoTa_060)
Hun von der Lippe hun hadde lært seg skikkelig.
she von der Lippe she had taught herself really
'That von der Lippe, she had really taught herself.'

In (4), the speaker refers to some person in charge of a club who, as we can guess from the context, did not grant her access to get in, perhaps because she had a fake ID card. The speaker expresses a distance to this person, she does not know her. In (5), the speaker talks admiringly about an actress, and the use of the demonstrative shows that the speaker needs to say that it is not a person close to her.

PDD Condition 2: The addressee does not personally know the person referred to.

- (6) Du vet han kjørelæreren jeg har?
you know he driving.teacher.DEF I have
'You know that driving teacher I have?' (NoTa_038)

⁴In both (4) and (5), the pronoun after the initial noun phrase shows that the initial noun phrase is left-dislocated. Left-dislocation is very common in speech, and is generally used to lift something into focus (Faarlund et al. 1997: 904–908).

- (7) Men **hun dattera** mi sa a syns det var så høyt under
 but she daughter.DEF my said she thought it was so high under
 taket.
 ceiling.DEF
 'But that daughter of mine said she thought it was so high under the
 ceiling.' (TAUS_a43)

In (6) the speaker is uncertain as to whether the addressee knows her driving instructor, and in (7) the speaker uses the PDD to express towards the hearer that she knows that the daughter referred to is not somebody the hearer knows.

PDD Condition 3: The speaker has a negative attitude to the person referred to.

- (8) Fordi han **idioten** vi hadde leid som sjåfør ...
 because he idiot we had hired as driver
 'Because that idiot we had hired as a driver ...' (NoTa_079)
- (9) Er ikke helt god hu mora mi altså ...
 is not quite good she mother.DEF my you.see
 'That mother of mine is not quite with it, actually ...' (Woman, web)

In (8) the speaker refers to a useless bus driver who has destroyed the engine of their bus by driving in second gear all the time, hence the use of the PDD (and of the word *idiot*). In (9) the speaker refers to a close family member (mother) who would presumably usually be referred to in a neutral or even loving way, but here there is a hint of dismay toward her, signalled by the PDD.

Other linguists have made similar observations on some of the conditions, such as Delsing (2003: 23), who says that the speaker is uncertain as to whether the listener knows who is intended (originally in Swedish: "talaren är osäker på om lyssnaren vet vem som avses"). The Swedish reference grammar (Teleman et al. 1999: vol. 2: 274) also points out that the listener's knowledge of the person referred to can be a condition of the use of this pronominal demonstrative. They say that the speaker and the listener must know of the person referred to, but that this person must not be a close acquaintance or be activated in the discourse.

More generally, both Condition 1 and Condition 2 have an aspect of politeness – what Brown & Levinson (1987) call positive politeness. Brown & Levinson (1987) suggest 15 politeness strategies divided into three supercategories: claim common ground, convey that S (speaker) and H (hearer) are cooperators, and

fulfil H's want for some X. I suggest that it is the cooperator strategy that is the closest to Conditions 1 and 2, and especially Strategy 9: *Assert or Presuppose S's knowledge of and concerns for H's wants* (Brown & Levinson 1987: 125). When the speaker does not know the referent personally, it is polite to convey this information to the hearer. Example (5) is a good illustration: If the speaker referred to the actress by her name but without the demonstrative, it would give the wrong impression to the hearer that the speaker actually knew the actress personally, or at least knew well the world of theatre and actresses in general. Using the PDD (by Condition 1) the speaker asserts that s/he wants the hearer to know that there is no such knowledge involved, in order not to mislead the hearer. Example (6) illustrates Condition 2, in which the speaker actively asks the hearer to direct her attention to the speaker's driving teacher, and in this way shows the hearer that she is aware of the hearer's need to know that this is not a person the speaker thinks the hearer knows already.

2.2 Demonstratives for fetching referents into the discourse (Lie 2010)

Lie (2008) focusses on the discourse activation function of the demonstratives under discussion, and says that their main task is that of bringing a new person into the discourse. Strahan (2007) also stresses the discourse activation aspect. Lie (2010) is more explicit, and says that we would get a better understanding of these demonstratives if we saw them in terms of background deixis, referring to what Diessel (1999) calls "discourse new and hearer old", i.e. that they do not refer to a previously mentioned referent or to a referent in the situational context, but that they are used to activate specific, shared knowledge (Lie 2010: 63).

Lie (2010: 63) also stresses the importance of grounding, "the shared knowledge, belief and attitudes of the interlocutors" (Croft & Cruse 2004: 60). A similar understanding to both Lie and Croft & Cruse can be found in the Swedish reference grammar (Teleman et al. 1999: vol. 2: 317): "the function of these words is to actualise referents that are not present in the concrete discourse, but who both speaker and hearer have in their world of concepts" [my translation, JBJ].

As we saw above, the Swedish reference grammar also stresses that the demonstrative must not refer to a close acquaintance, and it maintains that the demonstrative is "distance-creating" (Teleman et al. 1999: vol. 2: 317). Lie dismisses this possibility: "This I understand as an extra function, not as a primary function" [my translation, JBJ] (Lie 2010: 64). Lie also explicitly opposes Johannessen's criteria for the use of the PDD: "Her first two points say something about the fact that the speaker or hearer knows what is referred to, but not that the person referred to belongs in a common background. I therefore think it gives a bet-

ter understanding for these kinds of expressions that we here have background deixis” [my translation, JBJ] (Lie 2010: 70, fn. 15).

2.3 Accounts in the literature with relevance to the PDD

Himmelmann (1996) designs a typology of demonstratives in narrative discourse. The first type is situational use, which refers to entities in the utterance situation (1996: 219), the next is discourse deictic use, referring to events or propositions (1996: 224), and the third is the tracking use, to keep track of what is happening to whom (1996: 226). It is the fourth type, recognitional use (1996: 230), that is most interesting and relevant here, since this type appeals to knowledge of the referent. The intended referent, according to Himmelmann, is to be identified via specific, shared knowledge.

Diessel (1999) is work on which Lie (2010) is based. Diessel’s (1999: 106) system is developed from Himmelmann’s. He divides demonstratives into exophoric demonstratives (in the speech situation) and endophoric demonstratives, with three subtypes: anaphoric (referring to NPs), discourse deictic (referring to propositions) and recognitional (with the same name as in Himmelmann’s system). This latter type is the one that is relevant for Lie. They are used to activate specific shared knowledge; “[r]ecognitional demonstratives are specifically used to mark information that is *discourse new (i.e.) unactivated* and hearer old (...)” (Diessel 1999: 106, my italicization) and are “often used to *indicate emotional closeness*” (Diessel 1999: 107, my italicization), cf. also emotional deixis (Lakoff 1974).

It is clear that the demonstrative I am investigating here and the recognitional type have something in common. Indeed, Himmelmann (1996: 230) says that the speaker is uncertain as to whether the hearer can identify the intended referent, which he says can also be seen by the frequently accompanying tags like *you know?* or *remember?* As a matter of fact, (6) even contains this kind of appeal to the hearer, with a *you know*-question. At some level the PDD may appeal to the hearer’s knowledge, as in the obvious case of (6), *han kjørelæreren* (he driving.teacher.DEF) ‘that driving teacher’. Here the speaker may have talked about the driving teacher at some previous point in their common history. Using the demonstrative in addition informs the hearer that the person referred to is one that the speaker knows that the hearer does not know personally. But there are examples where it would be difficult to explain the use of the demonstrative as an appeal to common knowledge, as in (9), with *hu mora mi* (she mother.DEF my) ‘that mother of mine’. This phrase would be fully felicitous even if the hearer had never heard about the speaker’s mother before. Indeed, the demonstrative has no function here for the purpose of identification, and could have been dropped. Ev-

ery human being has a mother, and everybody knows this. Appealing to this kind of knowledge would undermine the whole idea of this function, since as humans we share a lot of common knowledge, which generally is not pointed out by grammatical markers. The demonstrative in (9) has only one function, which is that of conveying a negative attitude to the referent. Himmelmann's and Diessel's, and hence Lie's, appeal to shared knowledge does not work here.

Also, Himmelmann (1996: 210, 236) stresses that the recognitional demonstrative is unique and mentions the referent for the first time and above we saw the characterisation by Diessel (1999) (and Lie 2010): *discourse new and hearer old*. But not only does this not explain examples like (9), it is also descriptively incorrect. We will see many examples in §4 and §5 of the PDD being used repeatedly in short texts or discourses. Furthermore, the PDD can be used by both discourse participants for the same referent. This would also be difficult to explain if activation of common ground was the main purpose.

It seems that Himmelmann (1996), Diessel (1999) and Lie (2008; 2010) all see the PDD as a discourse structuring element, while Johannessen sees it as expressing subjectivity. Diessel (2006) sees such demonstratives as coordinating the interlocutors' joint focus of attention. Thus, unlike spatial demonstratives, the deictic centre is shifted from the physical world, i.e. the speaker's location at the time of the utterance, to a particular point in the unfolding discourse (Diessel 2006: 475). Johannessen's theory instead holds that there is a psychological space in which there is deixis, and where it is crucial to single out the psychological distance.⁵

Demonstratives that appeal to the speaker and/or hearer also appear in languages unrelated to the Scandinavian ones. Schapper & San Roque (2011) describe discourse demonstratives in Papuan languages. Evans et al. (2018) use the term "engagement" to refer to grammaticalised means for encoding the relative mental directedness of speaker and addressee towards an entity or state of affairs. These systems "express the speaker's assumptions about the degree to which their attention or knowledge is shared (or not shared) by the addressee" (Evans et al. 2018: 110). Clearly the PDD in the Scandinavian languages is part of an engagement system, with respect to shared knowledge of a referent. Condition 1 would be characterised as -SPKR.ENGAG (the speaker does not know the referent, and it is irrelevant whether the addressee does), while Condition 2 would be +SPKR-ADDR.ENGAG (the speaker knows the referent, but the addressee does not). Condition 3 would be +SPKR.ENGAG (the speaker knows the referent, and it is irrelevant if the addressee does).

⁵Proximal or neutral psychological distance are not expressed by a demonstrative, though it could be argued that the preproprial article, in those dialects that have it, has a proximal function (see Johannessen 2008: 170).

3 Method and material

In the following, I will investigate whether Lie (2010) is right, since he explicitly opposes Johannessen's analysis that the demonstratives are first and foremost used to show psychological distance. The way to do this, I suggest, is to look for discourses in which the PDD is used, and check empirically whether it has been used more than once in a discourse, or vice versa, whether there are contexts in which it ought to have been used as referent-invoking, but was not. If a text or discourse has multiple occurrences of the PDD for the same referent, it would indicate that Lie and also Teleman et al. (1999: vol. 2: 317) are wrong, and that the PDD does not have to be "discourse new". Furthermore, repeated use of the PDD to refer to the same person would probably be annoying for the listener, if indeed its function were to bring a referent into the discourse. If it does turn out that multiple uses of a PDD for the same referent can be found, it will also be central in demonstrating that the PDD expresses some psychological distance, as described by the three conditions in §2.1. Finally, if both discourse participants use the PDD for the same referent, it would be compatible with an explanation in terms of psychological distance (Condition 3), but not one in terms of invoking a referent from the common ground.

Finding appropriate examples of multiple PDDs has not been very easy. First, they are clearly a very oral phenomenon, regulating the emotional relationship between one or more of the interlocutors and other people mentioned. This means that written sources are only possible as long as they imitate spoken dialogue, such as in fiction. Even in fiction they are rare, because they are such an oral phenomenon; many authors probably do not think of them since they are not used to seeing them in writing. Transcriptions of spoken language are perhaps a better source, but only usable as long as they consist of dialogue, and of course, as long as the interlocutors speak about other people (while it may be possible to use the PDD with non-humans or non-animates, this has not been described in the literature). Second, both written and spoken language have in common the fact that whole noun phrases are not usually repeated over and over in a text or a discourse, since the speakers tend to use pronouns in these cases.

In spite of these challenges, there are some good sources. There are PDDs in two short stories by the acclaimed author Ingvild H. Rishøi. She has received numerous prizes for her work, including for her able use of the Oslo vernacular (*hverdagsspråket*). I have also found empirical evidence in two spoken language corpora. The NoTa-Oslo Corpus (Johannessen & Hagen 2008) is a corpus of 900,000 words of Oslo speech recorded in 2005, with the web-based user-

interface Glossa (Nøklestad et al. 2017), which makes it possible to search in the transcribed speech with audio and video presentations. It consists of carefully selected informants who together form a representative sample of the population, with respect to age, gender, educational and work background and location within Oslo, and contains semi-structured interviews and unstructured dialogues. The Nordic Dialect Corpus (NDC, Johannessen et al. 2009; 2014) has a Norwegian section consisting of 2.4 million words of spoken dialect dialogues. This corpus is also available with the Glossa search interface. Both corpora are available to researchers worldwide, and there is a password requirement giving most university employees the option to log in via their own university login system.⁶

Due to the challenges mentioned, it should be emphasised that there are not masses of examples of multiple PDDs. The examples that are found will each have to be studied and explained, and as we shall see, they provide evidence for determining which account is empirically correct. I should add that, as a native speaker, I find all the dialogues from the corpora and the language in the short stories fully natural.

4 Investigating two short stories

In this section, I will investigate two short stories by Ingvild H. Rishøi in order to find evidence for or against Lie's claim that it is background deixis that is the main function of the PDD, i.e. that it is used to activate specific background knowledge, fetching referents into the discourse. If this is its main function, we will not expect to find more than one PDD in a discourse.

4.1 *The life and death of Janis Joplin*

The short story *The life and death of Janis Joplin* (Rishøi 2014: 183–206) is about a 17-year-old girl. The girl, we understand, has problems with both school (she has reading difficulties, has a work practice placement, and wants to quit school permanently) and life. We follow her through her plans and her practice as a

⁶The data from the short stories and the corpora are referred to in the following way:

Rishøi 2014: 66, middle: A short story by Ingvild Rishøi, published in a volume in 2014, page 66, in the middle of the page

NoTa_027: From the NoTa corpus (Oslo), informant no. 027 (Johannessen & Hagen 2008)

TAUS_a43: From the TAUS corpus (Oslo), informant no. a43 (Hanssen et al. 1978)

NDC_kvam_04gk: From the Nordic Dialect Corpus, at the location Kvam, informant no. 04gk (Johannessen et al. 2009; 2014)

lumberjack, her love for her friend David, whom she met through her work practice, to the end, when she dies under a falling tree. The story is communicated through her thoughts and dialogues.

Central in the girl's story is an advisor at school, a man she often thinks of. He is referred to in her thoughts, and also with some cited dialogues. We understand from the story that he has made a deep impression on her. For example, his remark that she looks like Janis Joplin makes her want to start dressing like Janis Joplin and buy her CDs, which she listens to virtually all day and night, even as she is dying. Later in the story we learn that she despises this man (he has been unfaithful to his wife, whom he has forced to have an abortion, and probably worst of all, he has forgotten who our girl is). The story is written in an oral, lower register.

Since "discourse" is a central concept in the present chapter, it should be properly defined with respect to the data. This short story is written in a first person narrative (by the girl), and the story vacillates between the present time, which we can call speech time, when she is out in the forest cutting trees, and various episodes in the past at school, involving the advisor directly or indirectly, as well as episodes with other people, especially her friend David. Since the whole story is only 23 pages long, it could be seen as one long discourse (between the girl and us, the readers), but it is also possible to distinguish episodes, defined as sections that belong to the same time and space. In that case, the speech time would also be divided into different discourses. An alternative might be to regard the speech time as one long discourse, while the past episodes are each regarded as separate sub-discourses. The perspective taken is not arbitrary, since it will be decisive in regarding the discourse function of the demonstrative. However, the text itself guides us in how to divide the text into discourses: The girl's friend David is mentioned multiple times by name, both in the past episodes of the short story and the present speech time. This would have seemed very unnatural had the speech time been one continuous discourse, since it would then have been more natural just to refer to him by a pronoun. I shall therefore regard each stretch of text that belongs to the same time as one discourse. This gives us many small discourses, which will make the task of finding several PDDs with the same referent within one discourse more difficult, and hence also harder to find evidence for Johannessen's (2008) view.

There are 22 discourses, of which five involve the mentioning of the advisor, and 17 other discourses. The advisor is referred to six times by a non-pronominal noun phrase, five of them with a PDD, see (10)–(14).

- (10) ... men jeg satt på kontoret, og han rådgiveren løfta mappa mi og
but I sat in office.DEF and he advisor.DEF lifted file.DEF my and
pusta dypt.
breathed deeply
'... but I sat in the office, and that advisor lifted up my file and breathed
deeply.' (Rishøi 2014: 186, top)
- (11) ... jeg håper det er han fyren med bikkja som får de brevine.
I hope it is he guy.DEF with dog.DEF who gets those letters
'... I hope it is that guy with the dog that received those letters.' (Rishøi
2014: 187, bottom)
- (12) ... da han fyren kom til tredje vers.
when he guy.DEF came to third verse
'... when that guy got to the third verse.' (Rishøi 2014: 188, top)
- (13) Og da han ringte, han rådgiverfyren, så var det allerede for
and when he phoned, he advisor.guy.DEF then was it already too
seint.
late
'And when that advisor guy phoned, it was already too late.' (Rishøi 2014:
190, bottom)
- (14) Og etter tre måneder måtte jeg inn på rådgiverkontoret igjen, og
and after three months must I in on advisor's.office.DEF again, and
da var han fyren tilbake med skjegget sitt og hunden sin ...
then was he guy.DEF back with beard.DEF his and dog.DEF his
'And after three months I had to go back to the advisor's office again, and
that guy was back with his beard and his dog ...' (Rishøi 2014: 198, top)

The five occurrences of the advisor mentioned by the demonstrative plus a noun are each found in their own discourse, which, isolated, could mean that they were examples of the function of introducing or invoking a referent in the discourse. But if the demonstrative had this as its main function, we would expect the other protagonists, too, to be introduced by the demonstrative. This does not happen, however. There are many other people that are mentioned by name or by a noun phrase in the story, and they are never introduced by the demonstrative.

Some of the people referred to in the text are *sjefen* ‘the boss’ and *frøkena* ‘the teacher’, *kattedama* ‘the cat woman’ and *mamma* ‘mummy’. The advisor is also mentioned by his name, *Anders*, without a demonstrative. One important person who is close to the protagonist is referred to several times by a single name, *David* (and by pronouns); the girl obviously has no wish to distance herself from him. Also, others are mentioned, like *Janis Joplin* and *Leonard Cohen*. David is actually mentioned 25 times, and never introduced by a demonstrative, even though several of these times are in individual discourses, and not even close to each other: The first and second occurrences are in a discourse each on p. 187, but the third is not until p. 193, with eight discourses in between them.

If the main function of the PDD were to invoke a referent from the interlocutors’ common ground, we would expect this to be done for all the people who take part in this story. This does not happen. If only the important ones were to be invoked, i.e. those who are central in our girl’s story, we would expect not only the advisor, but also the narrator’s close friend David to be referred to by a PDD. This does not happen either. The main themes of the story are those that revolve around the advisor and David, but only one of them is referred to by the PDD. Clearly it is not fetching from the common ground that is the purpose of the PDD, since both characters should then have been introduced by it. What the PDD signals to us is instead that there is something not good about the advisor. It lets us understand that the narrator does not like him. Indeed, the author uses this demonstrative to make us understand what feelings the young girl has towards the advisor.

4.2 *Jimmy sa* ‘Jimmy said’

The short story *Jimmy sa* ‘Jimmy said’ (Rishøi 2014: 51–77) is about Jimmy, a young man who is looking back on his childhood and youth. It is told in the first person, mostly through spoken dialogue to a female workmate. We learn that his childhood was nice as long as his much older sister lived at home, and that they had a good relationship. She was blonde and they clearly looked different from each other, since he has black eyes, and is, we understand, dark-haired. He thinks he is adopted, but then he finds a photo of himself as a baby, with his sister. He confronts his mother with the idea that he is adopted, and she turns away. Then she confirms it, which leads him to start behaving violently and abusively. Later his sister comes home, and it turns out she is on drugs. Later again, his sister has moved into a flat with a man, Lars Arild, and when Jimmy visits them, she asks if he wants to move in with them, “as their son”, she explains. Jimmy runs out, and is upset for a long time. A couple of weeks later his sister dies of an overdose.

And some time later Jimmy realises the truth that his sister was his biological mother.

Although the story is short, it is useful to divide it into several discourses. The whole story is a monologue, with intermittent questions and answers to the workmate, and with subdialogues within the bigger monologue. I will count each part that belongs to the same temporal unit as a discourse, which gives 70 small discourse units. Again, this way of counting discourses makes the task of finding several PDDs in one discourse difficult, but it also would be a very convincing argument against the “hearer new” perspective.

Three people are referred to by the PDD several times. His teacher, Bente, is mentioned twice as *læreren min* ‘my teacher’ (without any demonstrative), and is then introduced by name, (15), after which she is referred to four times, either by name or by work title, but always with the PDD, (16)–(19). The first and the last occurrence, (16) and (19), are mentioned once each in a discourse unit, while (17)–(18) are in the same discourse.

- (15) Det var en ganske grei lærer egentlig, Bente het hu.
it was a quite nice teacher actually Bente called she
'It was a quite nice teacher, actually, Bente was her name.' (Rishøi 2014:
53, middle)
- (16) Og **hu** Bente fortalte at dem hadde blitt sjekka flere ganger, ...
and she Bente said that they had been checked several times
'And that Bente said that they had been checked several times ...' (Rishøi
2014: 54, top)
- (17) Men jeg snakka med **hu** læreren jeg fortalte om.
but I talked with she teacher.DEF I told about
'But I talked with that teacher I told you about.' (Rishøi 2014: 57, line 2)
- (18) Men **hu** Bente sto oppe ved tavla og fortalte om henne.
but she Bente stood up by blackboard.DEF and told about her
'But that Bente stood by the blackboard and told about her.' (Rishøi 2014:
57, line 5)

- (19) Så da sparka jeg ned alt mulig og fikk melding med
 so then kicked I down everything possible and got message with
 hjem hele tida, av **henne Bente** også.
 home all time.DEF from her Bente too.
 ‘So then I kicked down all sorts of things and got a letter to take home,
 from that Bente, too.’ (Rishøi 2014: 61, bottom)

It is clear that Jimmy knows his teacher, Bente, and also that he likes her, which he explicitly says the first time he introduces her by name, in (15). When the PDD is used with this person, it must be by Condition 2. Jimmy knows that the addressee does not know her, is polite and chooses to acknowledge this, and uses the PDD for the sake of the addressee. Since two occurrences, (17)–(18), are not only in the same discourse, but only two lines and two small sentences apart (and where both sentences also are about the teacher), Lie’s (2010) account cannot explain this use. Invoking a referent that is already activated does not make sense.

The father is referred to twice, both times with the demonstrative, seen in (20)–(21).

- (20) Jeg drømte om **han der faren** min i Romania.
 I dreamed about he there father.DEF my in Romania
 ‘I dreamed about that father of mine in Romania.’ (Rishøi 2014: 56,
 bottom)
- (21) ... og jeg så for meg **han der faren** min.
 and I saw for me he there father.DEF my
 ‘... and I imagined that father of mine.’ (Rishøi 2014: 65)

Jimmy has never met his father, does not know who he is, and knows nothing about him. He exists only in his imagination, he assumes his father is from Romania, presumably a country where people look like him. Here the PDD is strengthened by the distal adverb *der* ‘there’.⁷ Even though it is true that the addressee does not know his father, it is most likely the narrator’s own distant,

⁷The presence of the distal adverb *der* ‘there’ is not obligatory, and removing it would not change the truth conditions or the psychological distance of the phrase. Although the interplay between the PDD and the distal adverbs has not been studied in detail, I interpret these adverbs to further strengthen the distal meaning of the PDD. There is also a proximal adverb *her* ‘here’, which could plausibly be used, but it would not make the noun phrase psychologically proximal. There may be dialectal variation as to whether the proximal adverb can be used with PDDs. (For more on complex demonstratives, see Vindenes 2017.)

actually non-existent, relationship to the father that is in focus. It is a good example of Condition 1 of the PDD: The speaker does not personally know the person referred to.

Notice that the imaginary father is unique in the story. Jimmy only has one father, and there is no other father in the story. This is also the case for the mother and the sister, but in the case of the father, Jimmy chooses to make a point of the distance; that he does not actually know him, otherwise it might seem false, that he would try to make the father closer than he actually is. This explains the use of the PDD. Lie's account could also be used for the two occurrences of the father – not in the same discourse units – but would fall short of explaining why the other relatives and people in the story are not all invoked this way.

The most problematic figure for Jimmy in this story is his sister's cohabitant, Lars Arild. He is referred to by Jimmy six times, (22)–(27). The first three, (22)–(24), are mentioned once each in separate discourse units, but the last three are in the same discourse.

- (22) Jo, jeg hata han typen hennes skikkelig.
yes I hated he boy.friend.DEF her really
'Yes, I really hated that boyfriend of hers.' (Rishøi 2014: 65, bottom)
- (23) Nei, jeg var ikke der så ofte, på grunn av han fyren.
no I was not there so often on reason of he guy.DEF
'No, I wasn't there that often, because of that guy.' (Rishøi 2014: 66, middle)
- (24) Jo, jeg likte jo ikke han typen.
yes I liked yes not he guy.DEF
'Yes, I didn't like that guy, as you know.' (Rishøi 2014: 67, top)
- (25) "Lars Arild, kan du gi Jimmy litt saft?", og så henta han
Lars Arild can you give Jimmy some squash and then fetched he
typen saft ...
guy.DEF squash
"Lars Arild, can you give Jimmy some squash," and then that guy fetched
squash ...' (Rishøi 2014: 68, middle)

- (26) ... og så reiste han Lars Arild seg igjen.
 and then rose he Lars Arild himself again
 ‘... and then that Lars Arild rose again.’ (Rishøi 2014: 68, middle)
- (27) Og da blei søstera mi helt tårer i øya, og han Lars
 and then became sister.DEF my totally tears in eyes.DEF and he Lars
 Arild snudde seg.
 Arild turned himself
 ‘And then my sister became all tears in her eyes, and that Lars Arild
 turned.’ (Rishøi 2014: 69, top)

Lars Arild has been mentioned several times over these pages, and the PDD is used each time Jimmy mentions him. This is clearly because Jimmy does not like him, as he has said explicitly throughout the story. By using the PDD, he is able to indicate this distance every time. So although the PDD would not be strictly logically necessary, given that Jimmy has been explicit about his dislike of Lars Arild, it is used so that we are constantly reminded of his negative feelings.

It is central to point out that the last three occurrences are not only in the same discourse, but that in (25), the PDD is in fact in the same sentence as another mention of the same person, when Jimmy’s sister speaks directly to Lars Arild, immediately followed by Jimmy’s description of Lars Arild’s action.

Again, an account of the function of this demonstrative as one invoking a referent into the discourse cannot be right. The referent is already as foregrounded as it is possible to get.

Jimmy’s use of the PDD with Lars Arild is in sharp contrast to the one time where his sister mentions her boyfriend (28).

- (28) Jeg og Lars Arild har tenkt å gifte oss.
 I and Lars Arild have thought to marry ourselves
 ‘Lars Arild and I are planning to get married.’ (Rishøi 2014: 68, middle)

His sister can be assumed to love her boyfriend, and naturally does not use the PDD when she talks about him.

Finally, one non-central person is mentioned only in one discourse, but then with the PDD. This is the taxi driver who found him when he had run away from home (29).

- (29) Så han sjåføren hadde hørt at han skulle se etter en
so he driver.DEF had heard that he should look after a
trettenåring
thirteen.year.old
'So that taxi driver had heard that he should be on the lookout for a
thirteen-year-old ...' (Rishøi 2014: 70, middle)

The taxi driver is a peripheral character, and is only mentioned in one discourse. When the PDD is used here, it is due to Condition 1.

In this story, there are some people who are referred to without the PDD. These are people who we understand Jimmy thinks are nice and close to him. His sister, referred to by him as *søstera mi* 'my sister' or *Linn*, is referred to 50 times with a noun phrase, and never with a PDD, though there are of course many times that she is mentioned for the first time in a discourse, and sometimes with long stretches between, such as from the middle of p. 58 to the top of p. 60, with six discourses in between, dealing with aspects of his adoption. Another person who is never mentioned with the PDD is his mother, who is mentioned 30 times with a noun phrase: *mora vår* 'our mother', *mora mi* 'my mother', and *modern* 'mum', and also his work colleague *Freddy*.

There are three central characters referred to by the PDD, and they represent all three conditions of use. Jimmy's teacher, Bente, is referred to in accord with Condition 2: The addressee does not personally know the referent. This is done for politeness; Jimmy acknowledges that the addressee does not know her. His father is referred to by Condition 1: The speaker does not personally know the person referred to. This is also a function of politeness, but moreover, of honesty. Normally when people talk about their parents, there is an understanding between the interlocutors that there is a normal parent-child relationship involved. Since Jimmy neither knows his father nor anything about him, it would be almost dishonest to talk about him without the PDD, especially when the addressee is somebody he respects and wants to be honest to. The PDD is therefore necessary in this context. The sister's boyfriend Lars Arild is a person he dislikes intensely, as we have got to know throughout the story, both explicitly and implicitly by learning about Jimmy's behaviour when this person is mentioned. Every time the boyfriend is mentioned by Jimmy, it is with the PDD, and this is obviously in accord with Condition 3: The speaker has a negative attitude to the person referred to. Jimmy finds it necessary to distance himself from him every time he mentions him. Crucially, the sister does not use the PDD when referring to Lars Arild, who is a person she likes.

Interestingly, the PDD is used with different functions for the different people, but it is never difficult to understand which is applied with which person. On the contrary, the demonstrative helps us to interpret the relationship between Jimmy and the others.

If the function of the PDD were to fetch referents from the common ground, we would not expect some characters to be mentioned by it every time and others never (mother, sister, work colleague). Neither would we expect others to have to be fetched from the common ground only two lines apart (as with the teacher, (17)–(18)).

5 The use of the PDD in corpora of spoken conversations

Although the use of the PDD in the short stories in §4 shows quite clearly that its function is not to invoke referents from the common ground, it still remains to be seen whether this is also the case in genuine spoken language. To investigate this, it is useful to look at corpora of naturalistic, spoken dialogue. If the fictional dialogues in the short stories reflect the linguistic facts of the language, we should find equivalent patterns in the corpora, i.e. cases in which the PDD is used repeatedly with referents already invoked in the discourse. This is also what we find. We even find discourses where both interlocutors use the PDD for the same referent.

In dialect recording (30) an old woman from South Norway is interviewed about school, and the excerpt is from the beginning of a conversation about school times. The two occurrences of *hænn lerarn* (he teacher.DEF ‘the teacher’) are the only two mentioning the teacher.

(30) (NDC_kvam_04gk)

- a. [kvam_04gk:] visst e skulle leksa mi ælle no sjlikkt å **hænn lerarn** va strænng å hænn va nynåssjkær å de bøken våre va itte de
‘if I was going to read my homework or something, and **that teacher** was strict, and he was a Nynorsk person and those books of ours were not’
- b. [interviewer:] nei
‘no’
- c. [kvam_04gk:] næi menn ass åss hadde ei ænaste bok somm hæte “Såga omm fåLLke vårrt” å va på nynåssjk ællers så va de bokmål åss hadde å **hænn lerarn** hænn da skull e lessa et stykkji

‘no, but we had only one book that was called “The story of our people” that was in Nynorsk, otherwise it was Bokmål we had, and **that teacher** he was going to read a story’⁸

This excerpt shows that the woman uses the PDD the first time the teacher is mentioned, which could in isolation be interpreted as invoking him from the common ground; both interlocutors know that there is a teacher in a school situation. But then the PDD is used in the next sentence, too, making it very unlikely that there is invoking going on. Instead, the PDD can be understood by Condition 2: The referent is unknown to the addressee, and the speaker is politely acknowledging this fact by using the PDD.

In (31), an old woman (stamsund_04gk) and an old man (stamsund_03gm) from North Norway are talking about the old times when they played music in a school band, and the woman then introduces a Dane into the discourse:

- (31) (NDC_stamsund_04gk and NDC_stamsund_03gm)
- a. [stamsund_04gk:] va kje du mæ då **hann dannsken** va her å starpta opp
‘were you not there when **that Dane** was here and started up’
 - b. [stamsund_03gm:] jo va de ja e sjutta då hann e va mæ en par år
‘yes, I was, I stopped when he I was in it for a couple of years’
 - c. [stamsund_04gk:] ja e va mæ di to fôssjte åran ætte att **hann dannsken** starta opp jænn ætte de ha lie nere
‘yes, I was in it the first two years after **that Dane** started up again after it had been down’
 - d. [stamsund_03gm:] mm
‘mm’
 - e. [stamsund_04gk:] de va ennu hann farr din so kåmm inn i i leiliheta åt åss # å læmmpa inn en en kornætt å roppt “de herran de e te kjerringa n- må du bærre ha ho te å øv”
‘it was even your father who came into our flat and dumped a cornet and shouted “this one is for the woman, now you have to get her to practice”’
 - f. [stamsund_03gm:] mm
‘mm’

⁸Nynorsk and Bokmål are the two different written standards for Norwegian.

- g. [stamsund_04gk:] då bodd vi uttaførr
‘then we lived further away’
- h. [stamsund_03gm:] mm ja
‘mm yes’
- i. [stamsund_04gk:] å først att eg ha jo kje a me sjøl i de heile tatt å me
de va jo mått jo bynn å træn opp, menn dær uttaførr veit du i denn
gammle nolannsbangken jikk jo kje ann å håll på øve de va jo sækks
leilighete
‘and I didn’t choose it myself and I had to start practicing, but over
there you know in the old bank one couldn’t practice there as there
were six flats there’
- j. [stamsund_03gm:] nei **hann dansken** hann ee hann va då e va
‘no, that Dane he ehm he was’

Here, too, the PDD is used when the Dane is first introduced, but then the old woman continues to use the PDD in her next sentence as well, which would be surprising if the main function were to invoke the referent, who has just been invoked. The man also uses the PDD after three utterances by the woman. Again, the Dane is now well established in the discourse, so there is no need to further invoke him. Since they both use the PDD, it is evident that they do not invoke the referent for each other. Instead, the most likely interpretation is that both speakers for some reason distance themselves from this outsider from Denmark, and use the PDD (Condition 3) for this purpose.

In (32) we see an excerpt from a conversation between two 18-year-old men from Oslo. One introduces a new person, a German, into the discourse. Because of some laughter and self-interruptions there are some repetitions by this speaker. Crucially, every time the German (a cannibal) is mentioned, he is mentioned with the PDD both by the speaker and by his co-interlocutor.

- (32)
- a. [NoTa_016:] men hva med **han derre**, leste du om **han...** derre, **han tyskeren**, som hadde kuttet av utstyret på en fyr og spist det
‘but what about **that there**, did you read about **that there**, **that German**, who had cut off the equipment of a guy and eaten it?’
 - b. [NoTa_015:] ja ja ja
‘yes yes yes’
 - c. [NoTa_016:] **han derre kannibalen?**
‘**that there cannibal**’

- d. [NoTa_015:] han kannibalen ja
'that cannibal yes'

The German (the cannibal) is referred to with a PDD each time by both interlocutors. There is no other utterance in between the three times the PDD is used as part of a noun phrase. This would be very surprising if what they are doing is to invoke the referent from their common ground in every utterance and for each other. Instead, the use of the PDD here is to mark distance from themselves to this weird German (Condition 3). (For the function of the adverb *der* 'there', please see the comments in §4.2, on Jimmy's father.)

In (33), two girls, 18 and 19 years old, from the Oslo area, are talking about exams, and then turn to the topic of one of them, a short story. One of the girls starts discussing the main character in that story, using the PDD. In her next utterance she also uses the PDD about the same character.

- (33) a. [NoTa_093:] men den var så dum den novellen, **hu jenta** var så rar
'but it was so stupid that short story, **that girl** was so strange'
- b. [NoTa_094:] ja, den var så dårlig altså jeg klarte å skrive to og en halv side til slutt
'yes, it was so bad really, I managed to write two and a half pages in the end'
- c. [NoTa_093:] ja, **hu jenta** det handla om var teit
'yes, **that girl** it dealt with was stupid'
- d. [NoTa_094:] ja
'yes'

Since the PDD is used in two utterances in a row, there seems little reason to argue that invoking the referent is the function of this PDD. Instead, this use is in accordance with Condition 3: The speaker has a negative attitude to the person referred to. The speaker here clearly wants to distance herself from the referent, whom she characterises as strange and stupid.

In this section, we have investigated four short dialogues, i.e. four discourses, from real conversations, referring to a previous teacher, a Dane, a cannibal and a fictional character. In all of them, a PDD with the same referent was used several times in one discourse. In two of them, the PDD was used for the same referent by both interlocutors. In all cases the use of the PDD could be explained by reference to one of the three conditions in Johannessen (2008).

6 Summary and conclusion

Lie (2010) claims that the function of the demonstratives *hun* ‘she’ and *han* ‘he’ is to invoke a referent from the common ground of the interlocutors, following Diessel’s (1999) principle “discourse new and hearer old”, while Johannessen (2008) says that these demonstratives in Scandinavian are used to indicate psychological distance towards some other person referred to, hence the term “psychologically distant demonstratives”.

The aim of the present chapter was to test whether Lie’s analysis could be correct, and if not, whether Johannessen’s analysis would be. As empirical material, two short stories by the acclaimed author Ingvild Rishøi were chosen: *The life and death of Janis Joplin* and *Jimmy sa*. Both involve several people, of whom one or only a few are referred to by the PDD. In addition, four short dialogues from speech corpora containing spontaneous conversations were chosen (the NoTa-Oslo Corpus and the Nordic Dialect Corpus). Below I summarise the main findings. They all show that Lie (2010) cannot be right.

(i) The PDD are used for more than one mention per discourse: If the PDD were used to invoke a referent from a common ground, we would expect that referent to be mentioned at most once per discourse. Rather than regarding each short story as one discourse, I chose to divide them into several discourses, where each was a stretch of text coherent in space and time. This way of counting gave a high number of discourses per story (23 in the Janis Joplin story, and 70 in the Jimmy story), and therefore a low chance of finding the same character referred to by a PDD several times in one discourse. Even so, we did find this. In the story *Jimmy sa*, Jimmy’s teacher Bente is referred to twice (of four times in total) in the same discourse (three lines apart), while Lars Arild is referred to by him six times, with the PDD every time, three occurrences in the same discourse, and one of them even in the same sentence as another mention of the same referent.

In all four speech corpus dialogues, the PDD is used for the same referent, sometimes in the same sentence (the dialogue about the cannibal), or separated by one sentence (the dialogue about the Dane) or a one-word utterance (the dialogue about the teacher, and about the fictional character).

(ii) The PDD are used by different interlocutors within the same discourse: If the PDD had a purely discourse-regulating function, to bring a referent into the discourse from the common ground, we would not expect to see it used with one and the same referent by both interlocutors. However, in two of the naturalistic conversations, both interlocutors use the PDD (about the Dane and about the cannibal).

(iii) Not all referents are referred to by the PDD: If invoking referents into the discourse were the main function of the PDD, we would expect to find it with all characters who are not activated. However, this is not the case in the empirical material we have investigated here. Both the short stories contain many characters not referred to by the PDD: the girl's important friend David (25 times), and also her boss, mother, teacher, and somebody referred to as the cat woman. These characters are not all present in the discourse all the time, so they ought to have been fetched into the discourse, if this were the role of the PDD. Instead, they are possibly close to the main character or not central, and can be referred to without any particular marker. In the Jimmy story, there are also people who are not marked by the PDD: his sister (referred to more than fifty times), his mother and his workmate Freddy.

(iv) The PDD does not point out only key referents: It could be asked whether the PDD rather than pointing to referents of psychological distance instead points to referents that are central in the discourse. After all, both the girl's advisor and Jimmy's sister's boyfriend are central in the story. However, there are many others that are central or important, who are not introduced by a PDD. Jimmy's mother is in this category, as well as his sister. The girl's friend David is also very important, and presumably her mother. And vice versa, the taxi driver introduced by the PDD in Jimmy's story is not central. His teacher Bente is not very central either, but she does occur with the PDD. The function of the PDD, therefore, is not to point out the most central people.

Finally, one could ask whether Lie's function of invoking a referent from the interlocutors' background knowledge and Johannessen's PDD Conditions have a common core. Clearly, whenever a PDD is used for the first time in a discourse, it will invoke a referent. From the discussions above, though, this invoking is not the function of the PDD, because the PDD can be used several times in a short discourse (making the hearer quite annoyed in the end, if invoking were the central function), because both interlocutors can use the PDD about the same referent (and it would be comical if both interlocutors had to remind each other of the referent), and because not all referents are referred to by the PDD, even when they are not central in the discourse.

In conclusion, functions like background deixis and "discourse new and hearer old" do not cover the empirical facts. On the contrary, the referent could well be "discourse old", and even "hearer new", as when the speaker does not know the person referred to personally. The two short stories and four spoken dialogues investigated here show that the use of the PDD can instead be accounted for in terms of psychological distance (Johannessen 2008):

PDD Condition 1: The speaker does not personally know the person referred to.

PDD Condition 2: The addressee does not personally know the person referred to.

PDD Condition 3: The speaker has a negative attitude to the person referred to.

Acknowledgements

I would like to thank two anonymous reviewers and Yvonne Treis for their very constructive comments and suggestions that have greatly improved this paper.

Corpora

The Nordic Dialect Corpus: <http://www.tekstlab.uio.no/nota/scandiasyn/index.html>

The NoTa-Oslo Corpus: <http://www.tekstlab.uio.no/nota/oslo/english.html>

The TAUS Corpus of Oslo Speech: <http://www.tekstlab.uio.no/nota/taus/english.html>

References

- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Delsing, Lars-Olof. 2003. Syntaktisk variation i nordiska nominalfraser. In Øystein Alexander Vangsnes, Anders Holmberg & Lars-Olof Delsing (eds.), *Dialektsyntaktiska studier av den nordiska nominalfrasen*, 11–64. Oslo: Novus.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018. The grammar of engagement I: Framework and initial exemplification. *Language and Cognition* 10(1). 110–140.
- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Hanssen, Eskil, Thomas Hoel, Ernst Håkon Jahr, Olaug Rekdal & Geir Wiggen. 1978. *Oslo mål*. Oslo: Novus.

- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Johannessen, Janne Bondi. 2008. The pronominal psychological demonstrative in Scandinavian: Its syntax, semantics and pragmatics. *Nordic Journal of Linguistics* 31(2). 161–192.
- Johannessen, Janne Bondi & Piotr Garbacz. 2014. Proprial articles. *Nordic Atlas of Language Structures (NALS) Journal* 1. 10–17.
- Johannessen, Janne Bondi & Kristin Hagen. 2008. *Språk i Oslo. Ny forskning omkring talespråk*. Oslo: Novus.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli & Øystein Alexander Vangsnes. 2009. The Nordic Dialect Corpus – an advanced research tool. In Kristiina Jokinen & Eckhard Bick (eds.), *Proceedings of the 17th Nordic Conference on Computational Linguistics (NoDaLiDa 2009), May 14–16, 2009, Odense, Denmark*, 73–80. Northern European Association for Language Technology (NEALT).
- Johannessen, Janne Bondi, Øystein Alexander Vangsnes, Joel Priestley & Kristin Hagen. 2014. A multilingual speech corpus of North-Germanic languages. In Tommaso Raso & Heliana Mello (eds.), *Spoken corpora and linguistic studies*, 69–83. Amsterdam: Benjamins.
- Lakoff, Robin. 1974. Remarks on *this* and *that*. In Michael W. LaGaly, Robert A. Fox & Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, 345–356. Chicago, IL: Chicago Linguistic Society.
- Lie, Svein. 2008. Veldig sånn festejente. In Janne Bondi Johannessen & Kristin Hagen (eds.), *Språk i Oslo: Ny forskning omkring talespråk*, 78–95. Oslo: Novus.
- Lie, Svein. 2010. Om demonstrativer. *Maal og minne* 2. 59–78.
- Nøklestad, Anders, Kristin Hagen, Janne Bondi Johannessen, Michal Kosek & Joel Priestley. 2017. A modernised version of the Glossa corpus search system. In Jörg Tiedemann (ed.), *Proceedings of the 21st Nordic Conference on Computational Linguistics (NoDaLiDa 2017), May 23–24, 2017, Gothenburg, Sweden*, 251–254. Northern European Association for Language Technology (NEALT).
- Rishøi, Ingvild H. 2014. *Noveller i samling 2007–2014*. Oslo: Gyldendal Norsk Forlag.
- Schapper, Antoinette & Lila San Roque. 2011. Demonstratives and non-embedded nominalisations in three Papuan languages of the Timor-Alor-Pantar family. *Studies in Language* 35(2). 380–408.
- Strahan, Tania. 2007. The specific article in Norwegian. Reykjavík: University of Iceland. Unpublished ms.

9 Psychologically distal demonstratives in Scandinavian

Teleman, Ulf, Staffan Hellberg & Erik Andersson. 1999. *Svenska Akademiens grammatik*. Stockholm: Svenska Akademien.

Vindenes, Urd. 2017. *Complex demonstratives and cyclic change in Norwegian*. Oslo: University of Oslo. (PhD dissertation).

Chapter 10

Space, contrast and joint attention: Demonstrative adverbs in Russian, Estonian and Finnish

Tiina Nahkola

Maria Reile

Piia Taremaa

Renate Pajusalu

University of Tartu

This comparative study explores discourse functions of demonstrative adverbs in three areally close languages, which employ different demonstrative systems: Russian and Estonian (different two-term systems) and Finnish (an elaborate three-term system). We examine the use of demonstrative adverbs in a spatially contrastive setting using experimentally elicited data. We test whether the three chosen languages differ in terms of functions that demonstrative adverbs fulfil and whether the number of spatial distinctions within the demonstrative system affects the use and function of demonstrative adverbs in discourse reference. In all three languages, when referring to an object that can be conceptualised as a location, such as a building, the demonstrative adverbs are used in the following functions: i) identifying a referent, ii) tracking a referent, iii) conveying the information status of the referent. However, there are differences in how these languages use demonstrative adverbs to convey those functions. In the Russian and Finnish data, demonstrative adverbs are used mostly for tracking already activated referents, while in the Estonian data, demonstrative adverbs are a frequently used device for both identifying and tracking referents. In Finnish and Estonian, demonstrative adverbs can co-occur with demonstrative pronouns. These compound forms are used to indicate that the process of identifying the referent is unfinished. In all three languages, demonstrative adverbs are used both exophorically and anaphorically.



1 Introduction

Demonstratives (pronouns like *this* and *that*, and adverbs like *here* and *there*) are said to be a universal category (Diessel 1999; 2006; Dixon 2003), but languages differ remarkably as to the number and types of demonstratives they employ. A large number of studies have been conducted in an attempt to describe and categorise the different demonstrative systems found in the languages of the world (e.g. Diessel 1999; Kibrik 2011). A common basis for classification is the number and type of distance contrasts that are expressed with demonstratives (e.g. Anderson & Keenan 1985; Diessel 2013). In addition to the distance effect, there are less-studied aspects, such as the contrastive function (Meira & Terrill 2005), which is related to distance but is nevertheless a distinct function.

According to several accounts (Hanks 1992; 2011; Enfield 2003; Ariel 2013), the use of demonstratives in actual speech cannot be explained fully by the influence of distance and distance-related features. Instead, in these works, the emphasis has been on the exploration of conceptual access to the referents and different ways of directing attention. In addition, demonstratives can be viewed in terms of the discourse functions they serve (identifying or tracking a referent, marking definiteness, functioning as placeholders or pragmatic particles, etc.) (Himmelmann 1996; Diessel 1999; 2006). Nevertheless, there seems to be general agreement on the core functions of demonstratives – that is, to create and manipulate a joint focus of attention (Diessel 2006) – as well as on the strong association between spatial cognition and use of demonstratives (Coventry et al. 2008; Gudde et al. 2016).

Whatever the basis for the categorisation of demonstratives, it is important to note that demonstrative pronouns and demonstrative adverbs may encode different numbers of distinctions within a single system (see e.g. Hanks 2011). Despite operating within the same lexical category – namely, demonstratives – languages exhibit considerable variation in referential practices employed (Hanks 1990; Slobin 1996). The question then is: what is the relationship between the demonstrative system and the referential practices employed in a language? In other words, what is the role of system structure in discourse production (see also March & Pattison 2014)?

In the present comparative study, the focus lies on the use of demonstratives in a context that combines both spatial and anaphoric aspects of reference. In that, we study the functions of demonstratives in space and in discourse.¹ Whilst de-

¹We use the term “demonstrative” for both pronominal and adnominal forms, as well as for demonstrative adverbs (see Diessel 1999). When needed, we will specify what type of demonstrative is meant.

monstratives per se have attracted vast attention in both linguistic and philosophical research, cross-linguistic knowledge about them remains limited (Levinson 2018: 1). This is especially true for demonstrative adverbs, which have received less attention than demonstrative pronouns (see, however Laury 1996; 1997; Maes & Rooij 2007; Reile 2015; 2016; Reile et al. 2019). In this chapter, we observe the use of locative demonstrative adverbs (e.g. *siin* ‘here’ and *seal* ‘there’ in Estonian) in three languages: Finnish, Estonian and Russian. Finnish and Estonian are related languages; Russian is a contact language for both (moreover, we focus on the form of Russian spoken in Estonia). Finnish has an elaborate three-term demonstrative system, which contains hybrid forms of locative demonstratives that display both pronominal and adverbial features.² Estonian and Russian, in turn, display two different kinds of two-term demonstrative systems. Comparison of these three languages will enable us to assess how the demonstrative systems work in typologically different contact languages. More specifically, we seek to answer the following questions:

1. What functions do demonstrative adverbs serve in a spatially contrastive setting?
2. What is the relationship between the general demonstrative system of the language (the number and type of distance contrasts) and the way demonstratives are used for (discourse) reference?

First, we assume that the functions expressed with demonstrative adverbs in a spatially contrastive setting (a situation with competing referents) are not limited to indicating the location of the referent. Secondly, we assume that the properties of the general demonstrative system the language employs might affect the functions of demonstratives in discourse.

In this chapter, the focus is on referential expressions pointing to referents located in the space surrounding the interlocutors (see the methodology in §2). Following the traditional view on the pragmatic functions of demonstratives, the use of demonstratives occurring in our data is exophoric (Halliday & Hasan 1976).³ Exophoric reference cannot, however, always be distinguished from ana-

²Having the most complex demonstrative system among the languages in the sample, Finnish receives more attention than Estonian or Russian in this chapter.

³When used exophorically, demonstratives prototypically occur with gestures (Diessel 1999; 2006; Levinson 2004). However, there are also many uses of demonstratives that do not require gestures (Levinson et al. 2018). The data we use here does not allow us to observe systematically the gestures the speakers use. Therefore, we have decided to exclude this dimension from the present analysis.

phoric reference, since entities in the concrete surroundings of the speaker and the addressee can be referred to multiple times within discourse (see e.g. Levinson 2004). In fact, the simultaneous existence of multiple functions and interpretations is something characteristic of demonstratives.

In the following section, we will outline the research methodology. After that, in §3, we will give an overview of the demonstrative systems in Russian, Estonian and Finnish. §4 starts with a brief overview of the use of different referential devices in our data. This is followed by a closer examination of the use of demonstrative adverbs in each language. In the final section, we compare the results from all three languages.

2 Methodology

The data analysed in this study is a subset of the data from an experiment conducted to elicit referential utterances in a spatially contrastive context. It was a free production experiment in which participants had to use spoken language to describe and compare buildings that both the participant and the experimenter could see through a window.⁴ At the beginning of the experiment, the participant received written instructions pointing out the intended buildings (see Figure 1).



Instruction 1: there are two houses on the picture. Look out of the window and describe and compare the houses that have circles around them.

Instruction 2: now describe the house we are in and compare it one-by-one with the houses that have circles around them.

Figure 1: The two-sided instruction sheet: instructions given to the participants in Situation 1 (on the left) and Situation 2 (on the right). The numbers are illustrative and were not included on the original instruction sheet.

⁴For a more detailed description of the experiment, see Reile et al. (2019).

During the experiment, the experimenter stood next to the participant listening and giving minimal feedback. Resulting from the design of the task, the participants were operating with referents that were mutually known and identified – that is, presupposed (see Silverstein 1976) by both interlocutors. The data contains a total of 86 monologues collected from 25 Russian, 33 Estonian and 28 Finnish native speakers. When citing the data, we give each example an individual code, which consists of an indicator of the language, an identifier of the speaker and a number identifying the referential unit.⁵

The experiment consisted of two parts (Situation 1 and Situation 2). First, the participant described and compared two buildings, House 1 and House 2, of which House 1 was closer to the location of the participant and experimenter. In the second part, one more building (House 3) was added. This was the building where the participant and experimenter were located. By dividing the experiment into two parts, we were able to manipulate and observe the following factors: i) the change in the number of referents and thus the level of contrast, ii) the influence of distance and change in the deictic field. In addition to the aforementioned aspects, also other factors may influence the choice of the referential expression (for instance, the speaker's psychological distance from the referent). In this study, however, we concentrate on the effect of distance and contrast.

The data analysed in this study contains the referential units that were used in the first part of the experiment (henceforth called Situation 1) to refer to the building close to the speaker (House 1). In other words, out of two competing referents, the focus is on the closer one. Focusing only on the closer referent enables us to distinguish spatially contrastive use from other discourse functions of the demonstratives in the data. We can be certain that the use of distal demonstrative adverbs in referring to House 1 is not due to the comparison of House 1 to the closest referent (as could be the case in Situation 2, when comparing House 1 to House 3). The decision to focus only on a subset of the data results in an analysis based on a relatively small amount of data. Therefore, our approach in the present study is qualitative and the quantitative data presented is only intended to show the proportional use of demonstrative adverbs and to point out tendencies in their usage.

We have excluded from the analysis referential units that refer to only a part of a building. In some cases, the referential unit can refer to the location of the building or location inside the building instead of the building itself (see §4 for examples and discussion). In these cases, the speaker probably conceptualises the referent differently from those references which relate to concrete features of

⁵For example, E14.027. E = Estonian, F = Finnish, R = Russian.

the building (where the referent is conceptualised as an object). Nevertheless, we have included in our data also those referential units that refer to the referent's location.

The experiments in all three languages were conducted and video-recorded at the same location in Tartu, Estonia. The recorded data was manually transcribed and annotated by native speakers. The transcription system employed was the Jefferson system (Jefferson 2004). The data was annotated for a number of referential devices (see also Reile et al. 2019). For the present chapter, the most relevant categories are the first two mentioned and the last on the list: demonstrative pronouns (in pronominal and adnominal use; later called BareDem and DemNP), demonstrative adverbs (DemAdv), NPs occurring without a demonstrative (BareNP), third person pronouns (PersPron), zero reference and different combinations of demonstrative pronouns, demonstrative adverbs and NPs.

One more factor taken into account during the annotation of the data was the number of consecutive utterances in which the same referent was mentioned (that is, the tracking function of the referential unit). This factor, which henceforth is referred to as "mention number", is related to the (dis)continuity of the referential chains formed by multiple references to the same entity. It is important to note that in this chapter, the term "first mention of the referent" does not indicate the absolute first time the referent is mentioned in the current discourse, but instead the first mention of the referent within the current referential chain. The annotation of the mention number enables us to explore the association between the information status of the referent and the choice of the referential expression. Following the terminology initiated by Gundel et al. (1993; 2010), we consider the first mention of the referential chain to comply with the status "familiar". Subsequent mentions, in turn, comply with the statuses "in focus" and "activated".

3 Demonstratives in Russian, Estonian and Finnish

Estonian and Finnish as Finnic languages and Russian as a Slavic language have a three-directional system for demonstratives (goal, source and location). None of the languages makes a formal distinction between the demonstrative pronouns in their pronominal use and adnominal use (that is, as demonstrative determiners), but all three languages make a formal distinction between demonstrative pronouns and demonstrative adverbs. It should be noted, however, that in Finnish, the local cases of the demonstrative pronouns share some morphological, syntactic, and semantic features with the demonstrative adverbs. Therefore, the dis-

tinction between a demonstrative pronoun and an adverb is sometimes difficult to make in Finnish (see §3.3 and Laury 1996). The demonstratives used for establishing a new referent are not formally distinguished from those that indicate a contrast between two already established referents. In Finnish and Estonian, demonstrative adverbs can be used adnominally, whereas in Russian this kind of use is not typically found. Estonian and Russian employ distance-based two-term systems; Finnish, in turn, has a mixed person- and distance-based system with three terms.

In the following sections, we present the most central features of the demonstrative systems in Russian, Estonian and Finnish.

3.1 Russian demonstratives

There are two demonstrative pronouns – proximal *eto* ‘this’ and distal *to* ‘that’ – in contemporary Russian (Sheljakin 2002: 118; Timberlake 2004: 233). The spatial function of demonstratives activates mostly when two referents are in contrast; in other contexts, discourse factors are more important (Grenoble 1998). Both have the same morphological stem (*to*). In older Russian, the demonstrative *se* was used for proximal reference, and this stem has been preserved in demonstrative adverbs up to the present.

Demonstrative adverbs are as follows: for location, *tut* ‘here’, *zdes* ‘here’ and *tam* ‘there’; for goal, *sjuda* ‘to here’, *tuda* ‘to there’; for source, *otsjuda* ‘from here’, *ottuda* ‘from there’. Russian demonstrative adverbs do not normally co-occur with a noun phrase. There are some examples like *tam, v nebesah* ‘there in the sky’ in written Russian, but the construction is better analysed as a case of apposition (see Sakhai 2003) than as a single phrase that contains a demonstrative adverb.

Tut ‘here’ and *zdes* ‘here’ are both proximal demonstrative adverbs. Grenoble (1998: 105–106) states that *zdes* indexes a concrete location – a physical space with dimension (*Godami on rabotal zdes*). ‘For years he worked here [in the office]’). *Zdes* is used in opposition to *tam* ‘there’. Moreover, when the ground is a concrete place, only *zdes* is acceptable. *Tut* may also index a place when the place is an abstract space. The underlying opposition of ‘here’ to ‘there’, which is inherent to *zdes*, distinguishes it from *tut*, which does not signal this opposition. Rather, *tut* refers to a place with undetermined or irrelevant boundaries (*Pochemu ne poznakomilsja tut ni s odnoj iz nih?* ‘Why haven’t you met a single one here [undefined space]?).

3.2 Estonian demonstratives

Estonian has two demonstrative pronouns: *see* ‘this’ and *too* ‘that’. *Too* has regionally varying use. In northern Estonia, *too* is not used in spatial reference. In anaphoric reference, its use tends to be restricted to time expressions and in reference to the second human character of a literary narrative (Pajusalu 2006). Therefore, based on spatial distinctions made by demonstrative pronouns, Estonian has two demonstrative pronoun systems: a one-term system, where *see* ‘this/that’ is distance-neutral (Reile 2015; Pajusalu 2009) and a two-term system, where *see* ‘this’ is the proximal and *too* ‘that’ the distal demonstrative pronoun (Reile 2016). In addition, both demonstrative pronouns can be used as definite determiners but *see* is used much more frequently in this function (Pajusalu 2009). Moreover, the demonstrative pronoun *see* and third person pronoun *tema/ta* share the anaphoric referential domain. In addition, both of these pronouns can be used to refer to animate and inanimate referents depending on their activation status (Pajusalu 2017).

Estonian has six demonstrative adverbs displaying two distance-based contrasts, goal-based *sia* ‘hither’ and *sinna* ‘thither’, location-based *siin* ‘here’ and *seal* ‘there’, and source-based *siit* ‘from here’ and *sealt* ‘from there’. In spatial reference, *siin* ‘here’ is the proximal and *seal* ‘there’ is the distal demonstrative adverb. Demonstrative adverbs can also be used as definite determiners in place denoting NPs (e.g. *seal majas* ‘(in) yonder house’). Moreover, distal demonstrative adverbs can be used to denote previously mentioned location in a narrative (Pajusalu 2017).

3.3 Finnish demonstratives

The Finnish language has three demonstrative pronouns: *tämä/tää* ‘this’, *tuo/toi* ‘that’ and *se* ‘it, that, the’.⁶ According to the traditional, largely distance-based view (e.g. Larjavaara 1990), *tämä* ‘this’ refers to a referent near the speaker and *tuo* ‘that’ to a referent distant from both the speaker and the addressee. This view is not entirely based on distance, as it considers *se* ‘it, that, the’ to be distance-neutral. Moreover, *se* is hearer-centred, whereas *tämä* and *tuo* are speaker-centred. However, several accounts (Itkonen 1966; Laury 1997; Seppänen 1998; Etelämäki 2006) have proposed an analysis relying more on the social and interactional functions of the Finnish demonstratives.

In addition to its hearer-centred meaning, *se* functions as an anaphoric device, maintaining reference when the referent is already activated and sufficiently de-

⁶ *Tämä* and *tuo* are used in standard Finnish, *tää* and *toi* in many colloquial varieties.

fined (Laury 1997; Etelämäki 2005; Duvallon 2005). *Se* can also identify a referent that is considered to be known in advance (Laury 1997). *Tämä* and *tuo* are primarily exophoric (Hakulinen et al. 2004), but can also be used anaphorically. Anaphoric *tämä* and *tuo* indicate a higher level of salience for the referent than *se*. Moreover, *tämä* and *tuo* indicate open reference – in other words, they imply that the process of identifying the referent is still ongoing. *Se*, on the other hand, marks closed reference (Etelämäki 2005). *Tuo* has also been reported to show a speaker's dissociation or uncertainty in relation to a referent (Hakulinen 1985; Laury 1997).

Finnish demonstrative adverbs share the same three-stem system with demonstrative pronouns. Demonstrative pronouns inflect for most Finnish cases, including for six cases that indicate a place of action. These local cases are divided into two groups according to whether the action takes place within a three-dimensional location (internal cases) or outside or on a two-dimensional location (external cases) (Hakulinen et al. 2004). The locative demonstrative adverbs have only three forms: location, source and goal. Only pronouns have plural forms, but adverbs can co-occur with a noun in plural form. The paradigms of the local case forms of the pronouns and the demonstrative adverbs partially overlap. All (singular) forms are presented in Table 1.⁷

Table 1: Spatial demonstratives in Finnish (location, source, goal)

DEM stems	Adverbs	Internal case forms of the pronouns	External case forms of the pronouns
<i>tä-</i>	<i>täällä, täältä, tänne</i>	<i>tässä, tästä, tähän</i>	<i>tällä, tältä, tälle</i>
<i>t(u)o-</i>	<i>tuolla, tuolta, t(u)onne</i>	<i>t(u)ossa, t(u)osta, t(u)ohon</i>	<i>t(u)olla, t(u)olta, t(u)olle</i>
<i>se-</i>	<i>siellä, sieltä, sinne</i>	<i>siinä, siitä, siihen</i>	<i>sillä, siltä, sille</i>

There is not full agreement on the categorisation of the Finnish spatial demonstratives (see Laury 1996 for an overview of the discussion). Laury (1996) suggests treating the local demonstratives as a continuum, where the adverbs are the most adverbial and the external case forms the least adverbial forms. The label “the

⁷The brackets indicate the parts that are often omitted in spoken language, for example *tuonne* > *tonne*. In addition, the suffixal parts of the demonstratives are often shortened, e.g. *täällä* > *täääl*, *siinä* > *siin*.

least adverbial” refers to the syntactic and semantic features of the external cases: in actual speech, they are primarily used to express non-locative functions – for example, possessor or recipient – and only occasionally refer to locations (Laury 1996; 1997). In addition, they always behave syntactically like pronouns. The internal local cases, in turn, can manifest an adverb-like behaviour. In this study, the term “internal local cases” refers to these locative forms of the pronouns that carry also adverb-like features.

Both demonstrative adverbs and demonstrative pronouns can be used adnominal (Hakulinen et al. 2004). When occurring with a noun, the pronouns tend to agree with the head of the noun phrase in both case and number, whereas the adverbs do not. Similar to adverbs, the internal case forms of demonstrative pronouns do not always agree with the head of the noun phrase. Moreover, the internal cases can co-occur with an adverb or a postpositional phrase, unlike other pronouns. How then do speakers make choices between pronominal and adverbial demonstratives in spatial reference in Finnish? The following factors, among others, have been suggested to affect the choice between the forms: exactness of reference, visibility of the referent, proximity of the referent and size or boundedness of the referenced area (e.g. Itkonen 1966; Siitonens 1979; Östman 1995).⁸

Even though the internal local case forms of pronouns resemble adverbs in many ways, the internal cases are not always used to indicate location. Like the external case forms, the internal cases can express non-local concepts such as a part–whole relationship or physical/mental state. Adverbs, on the other hand, do not express any concepts other than concrete location (at least in our data).

4 Results

In the present analysis, the focus is on referential units referring to the proximal building (House 1) during the first part of the experiment (Situation 1). The focus of the analysis is on demonstrative adverbs, but we will begin by giving an overview of all referential devices that are used to refer to House 1 in Situation 1 (see Table 2).

The most striking differences between the data for the three languages are the following: i) the proportion of referential units that include a demonstrative is

⁸In addition to the authors cited, Laury (1996) proposes an interesting analysis based on the conceptualisation of scenes in terms of figure and ground (see also Talmy 1983). The referents referred to by internal cases tend to be conceptualised as figures and those referred to by adverbs as grounds.

Table 2: All referential units referring to House 1 in Situation 1 (Bare-Dem = demonstrative (pronoun or adverb) without a separate nominal head; DemNP = demonstrative (pronoun or adverb) with a separate nominal head; BareNP = noun phrase without a demonstrative; Pers-Pron = third person pronoun; Zero = zero reference)

Language	Referential device					Total
	BareDem	DemNP	BareNP	PersPron	Zero	
RUS	47 (17%)	29 (10.5%)	113 (40.9%)	68 (24.6%)	19 (6.9%)	276 (100%)
EST	74 (18.6%)	89 (22.4%)	177 (44.6%)	45 (11.3%)	12 (3%)	397 (100%)
FIN	218 (57.1%)	113 (29.6%)	45 (11.8%)	—	6 (1.6%)	382 (100%)

larger in the Finnish data than in the data for the other two languages; this is especially true for the bare demonstratives (that is, demonstratives without a separate nominal head); ii) conversely, Finnish has the smallest number of referential units that do not include a demonstrative (BareNP). Moreover, Finnish does not employ personal pronouns for inanimate referents, whereas in the Russian data personal pronouns are the second most frequent referential device. In most of the categories (nominal and adnominal demonstratives, personal pronouns and zero reference) presented in Table 2, Estonian is situated between Finnish and Russian when it comes to the frequency of different referential devices. The proportion of the referential units made up of bare NPs is, however, the largest in the Estonian data.

Next, we will describe the use of demonstrative adverbs in the three languages (see Table 3).

Table 3: Demonstrative adverbs referring to House 1 in Situation 1

Language	Demonstrative adverb			Total
	Proximal	Distal	Hearer-centred/Anaphoric	
RUS	15 (44.1%)	19 (55.9%)	—	34 (100%)
EST	33 (62.3%)	20 (37.7%)	—	53 (100%)
FIN	1 (5.0%)	4 (20.0%)	15 (75.0%)	20 (100%)

Although Finnish has the greatest number of referential units with a demonstrative, it has the smallest number of units with a demonstrative adverb.⁹ The Estonian data includes more demonstrative adverbs than the Russian data, which is in line with the overall frequency of demonstratives in these languages in our data (presented in Table 2). In the Russian data, distal adverb forms are the most frequent, unlike in the Estonian data, where proximal forms outnumber distal forms. The Finnish data includes mostly hearer-centred/anaphoric demonstrative adverbs. In this context, only the anaphoric function is relevant, as the referent is at the same distance from both interlocutors. The hearer-centred function activates only in situations where the addressee's current sphere is conceived as distinct from the speaker's sphere (see Larjavaara 1985).

We will discuss the data for each language more thoroughly in the following sections.

4.1 The use of demonstrative adverbs in the Russian data

In the Russian data for House 1 in Situation 1, there are a total of 34 referential units with demonstrative adverbs, 15 proximal (14 *zdes'* 'here', 1 *tut* 'here') and 19 distal (*tam* 'there') (see Table 4). In the Russian data, demonstrative adverbs do not function as determiners.

Table 4: Use of Russian demonstrative adverbs in referring to House 1 in Situation 1

Adverbs	1 st mention	2 nd mention	3 rd mention and further	Total
tut/zdes'	9 (60.0%)	5 (33.3%)	1 (6.7%)	15 (100%)
tam	—	7 (36.8%)	12 (63.2%)	19 (100%)

Zdes' and *tut* always refer to House 1 in Situation 1, see (1) and (2). Distal *tam* refers mostly to House 2, but sometimes also to House 1, see (3). This, in addition to the fact that *tam* is not used for first mentions in the referential chain, suggests that in addition to indicating distance, *tam* can function as an anaphoric device. When *tam* is used anaphorically, the contrast between the referents is not relevant.

⁹It should be noted that also the internal local cases of Finnish demonstrative pronouns exhibit adverb-like behaviour. These forms are not included in Table 3. Instead, we discuss the use of the internal local cases in §4.3.

(1) (R20.041)¹⁰

а (.) трубы (.) по которым стекает вода, здесь я их вижу на каждом уголку, они стоят (.) висят. а там (.) там я вообще их не вижу.

a, truby, po kotorym stekaet voda, zdes' ja ih
o pipe.PL PREP REL flow.down.3SG water.NOM here 1SG 3PL.ACC
vizhu na kazhdom ugolku, oni stojat, visjat. a tam,
see.1SG PREP every.LOC corner.DIM.LOC 3PL stand.3PL hang.3PL but there
tam ja voobsche ih ne vizhu.
there 1SG at.all 3PL.ACC NEG see.1SG

'O, pipes in which water flows down, here I can see them on every corner,
they are standing, hanging. But there, there I can't see them at all.'

(2) (R06.061)

... тут написано ресторан-бар.

... tut napisano restoran bar.
here.DEM.ADV.LOC write.PERF.PTCP restaurant.bar.NOM

'Here it says "restaurant-bar".'

(3) (R22.007)¹¹

так же я думаю там два входа,

tak zhe ja dumaju tam dva vhoda,
also PRTCL 1SG think.1SG there.DEM.ADV.LOC two entrance.GEN

'Also, I think there are two entrances there.'

While the usage contexts of *tam* are mostly different from the contexts typical for the other two demonstrative adverbs, it is more difficult to pinpoint the differences between the adverbs *zdes'* and *tut*. However, some tendencies can be detected. In this part of the analysis, we take into account the data concerning both House 1 and House 3 (the building where the interlocutors are located). This enables us to provide a more comprehensive account of the usage contexts of the two proximal adverbs. First, 83% (n=10) of the occurrences of *tut* refer to House 3, whereas with *zdes'*, the proportion of the units referring to House 3 is smaller (64%, n=27). Consequently, *zdes'* is used to refer to House 1 more often than *tut*.

¹⁰See the appendix for a list of symbols used in the transcription.

¹¹The preceding context for the example is the following: вот, в принципе здание небольшое. эх, что ещё можно сказать? много окон. так же я думаю там два входа, если я не ошибаюсь. 'So practically it is not a big building. Oh, what else could I say? Lots of windows. I also think there are two entrances, if I am not wrong.'

(*zdes'* 36%, n=15; *tut* 17%, n=2). Secondly, the analysis of the mention numbers for the referential units reveals another difference between the usage patterns of these two adverbs. 75% (n=9) of the referential units containing *tut* are the first mentions of their referential chain, whereas with *zdes'*, the proportion of first mentions is 48% (n=20). To conclude, our data suggests that the main function of *zdes'* is contrastive, and *zdes'* is less often used for identifying a referent at first mention than *tut*. In our context, *tut* seems to specialise for House 3, which is the location of the interlocutors. As the boundaries of House 3 are not observable for the speaker, this building is conceptualised as a space, while House 1, which can be seen through a window, is conceptualised as an object. This supports Grenoble's (1998) finding discussed in §3.1: *zdes'* indexes a concrete location and is used in opposition to *tam* 'there', while *tut* refers to a place with undetermined or irrelevant boundaries.

4.2 The use of demonstrative adverbs in the Estonian data

In the Estonian data for House 1 in Situation 1, *siin* 'here' is used 33 times and *seal* 'there' 20 times. The analysis of the mention numbers of the referential units reveals that the use of *seal* increases and *siin* decreases as the mention number increases (see Table 5). This suggests that in our data, the majority of the referential units with *siin* are used spatially (4), whereas *seal* is used mostly anaphorically (5). However, anaphoric use of *siin* is also possible (6).

(4) (E15.003)

sellel	majal	siin	on	(.)
this.DEM.PRON.ADE	house.ADE	here.DEM.ADV.LOC	be.PRS.3SG	
selline	uhke	silmatorkav	katus	...
this.kind.NOM	grand.NOM	conspicuous.NOM	roof.NOM	

'This house here has a grand roof ...'

Table 5: Use of Estonian demonstrative adverbs in referring to House 1 in Situation 1

Demonstrative adverb	1 st mention	2 nd mention	3 rd mention and further	Total
<i>siin</i>	19 (57.6%)	11 (33.3%)	3 (9.1%)	33 (100%)
<i>seal</i>	3 (15.0%)	3 (15.0%)	14 (70.0%)	20 (100%)

- (5) (E17.001; E17.003; E17.004)

nii **esimesel majal** on selline pruunikas punane so first.ADE house.ADE be.PRS.3SG this.kind.NOM brown.NOM red.NOM katus? (.) üsna vana **maja** tundub olema, ilmselt roof.NOM rather old.NOM house.NOM seem.PRS.3SG be.INF apparently on seal olnud mingi restoran või be.PRS.3SG there.DEM.ADV.LOC be.PST.PTCP some.kind restaurant.NOM or baar, bar.NOM

'(The) first house has a brownish red roof. (It) looks like quite an old house, there has probably been a restaurant or bar.'

- (6) (E01.051; E01.052)

esimene **maja** on väljast väga hästi valgustatud first.NOM house.NOM be.PRS.3SG outside.ELA very well light.PST.PTCP **siin** on lausa päris mitu valgustit ees ... here.DEM.ADV.LOC be.PRS.3SG even quite many lamp.PART in.front
'The first house is very well lit, there are many lights here ...'

When the speaker refers to House 1 in Situation 1 with *seal*, the same referent has usually been mentioned in the previous utterance and is in the current focus of attention (5).

Estonian demonstrative adverbs can be combined with NPs and demonstrative pronouns (4). While in general *see* 'this' can occur with both distal *seal* and proximal *siin*, *too* 'that' seems to be combined only with distal *seal* (e.g. Reile 2016). However, in the current data, *see* occurs only with proximal *siin*.¹² The compound forms are used mostly, but not always, for the first mention of the referent in the referential chain.

The adverb *siin* usually comes after the pronoun *see* (and possible NP) in the utterance (7).

- (7) (E30.031; E30.033)

siin ei tea mis seal katuse here.DEM.ADV.LOC NEG know.PRS.1SG what there.DEM.ADV.LOC roof.GEN all on (...) see **siin** jah?
under be.PRS.3SG this.DEM.PRON.NOM here.DEM.ADV.LOC yes

¹²There were no instances of *too* used in reference to House 1 and therefore also no combinations with the adverbs.

on tänavale ka rohkem avatud (.) kuna autod
be.PRS.3SG street.ALL also more open since car.PL.NOM
sõidavad mööda ...
drive.PRS.3PL by

‘Here I don’t know what is under the roof (up) there. This here is more open to the street as well since the cars drive by ...’

Siin can also occur before *see* (8), but the occurrence of *siin* in this position is very rare in our data (only one use for House 1 in Situation 1).

- (8) (E31.020)

sellel tagumisel on katuseaknad aga
this.DEM.PRON.ADE hindmost.ADE be.PRS.3SG skylight.NOM.PL but
siin sellel esimesel majal ei ole
here.DEM.ADV.LOC this.DEM.PRON.ADE first.ADE house.ADE NEG be
katuseaknaid,
skylight.PART.PL

‘This one at the back has skylights but here this house in the front does not have skylights.’

In (8), where *siin* occurs before *see*, the referential unit is the first mention of the referent in the current referential chain and the referent is not in focus. The speaker uses *siin* to define the visual region in which the addressee should look for the referent, and then identifies the house by using a demonstrative pronoun and a lexical NP (*sellel esimesel majal*). In (7) where *siin* comes after *see*, the house has been mentioned in the previous utterance and is in the current focus of attention. This pattern of demonstrative adverb use suggests that in Estonian, the position of demonstrative adverbs in a referential utterance can be used to connect the visual world with the cognitive status of the referent and through this to indicate whether or not the referent is in the current joint focus of attention of the interlocutors. This is in line with previous findings regarding the use of Estonian demonstrative adverbs for spatial reference (Reile 2015; 2016).

In our data, when referring to the buildings, Estonian speakers also use demonstrative pronouns and third person pronouns – which are common anaphoric devices in Estonian (Pajusalu 2009). In Estonian, both demonstrative adverbs and demonstrative pronouns occur as determiners in place-denoting NPs. While demonstrative pronouns and third person pronouns can occur in any case form, internal case forms are central to the current study. More specifically, the case form that is most of interest is the inessive (an internal case form), e.g. *selle-s*

maja-s ‘in this/that house’, where *-s* is the inessive case marker. The alternative referential devices that could be used instead of a pronoun in the inessive case are the demonstrative adverbs. However, since demonstrative adverbs can occur with demonstrative pronouns, these referential devices can also complement each other, instead of being only alternative devices.

To conclude, the proximal demonstrative adverb *siin* is used mainly spatially and the distal demonstrative adverb *seal* is used anaphorically when referring to House 1 in Situation 1. The longer the referential chain, the more instances of distal *seal* and the fewer of proximal *siin* occur. Of the two demonstrative adverbs in Estonian, only *siin* occurs in demonstrative pronoun and adverb combinations in our data. By combining demonstrative pronouns and demonstrative adverbs in the same referential unit, the speaker can provide information on the referent at different levels of discourse as well as double the force of the spatial reference.

4.3 The use of demonstrative adverbs in the Finnish data

In the Finnish data, demonstrative adverbs are used for referring to House 1 twenty times altogether in Situation 1 (see Table 6). If we also consider the use of the internal case forms of demonstrative pronouns (presented later in Table 7), this number increases to 102. In the present analysis, the internal local cases of the Finnish demonstrative pronouns are taken into account when they appear without a head noun (9).

(9) (F15.014)

tä^s on tietysti vaan (.) ravintola ja
this.DEM.PRON.INE be.PRS.3SG of.course only restaurant.NOM and
baari?
bar.NOM

‘Here is, of course, just a restaurant and a bar.’

Table 6: Use of Finnish demonstrative adverbs in referring to House 1 in Situation 1

DemAdv stem	1 st mention	2 nd mention	3 rd mention and further	Total
tää-	–	–	1 (100%)	1 (100%)
t(u)o-	1 (25.0%)	–	3 (75.0%)	4 (100%)
se-	–	4 (26.7%)	11 (73.3%)	15 (100%)

The internal local cases of pronouns are analysed separately from the adverbs in view of their somewhat different usage patterns. Therefore, we have excluded from the present analysis those forms which cannot be unambiguously placed in either the adverb or pronoun category. This means that internal case forms occurring in the same case as a head noun, as in (10), are excluded.¹³

- (10) (F28.051)

tässä (.) lähemmässä talossa on myös
this.DEM.PRON.INE near.COMP.INE house.INE be.PRS.3SG also
tollaset isot (.) valot?
such.NOM.PL big.NOM.PL light.NOM.PL

‘In this nearer house, there are also those big lights.’

The external case forms of the demonstrative pronouns are not included in the analysis, since their adverb-like use is not found in our data and their frequency is low.¹⁴

In our discussion, first, we focus on the demonstrative adverbs only. Then, we discuss the use of the internal local cases in comparison with the use of the demonstrative adverbs.

The speakers refer to House 1 with demonstrative adverbs mostly when the referent has already been mentioned in one or more previous utterances. The most frequent stem is the hearer-centred/anaphoric *se*-stem.¹⁵ This is in line with the previous findings showing that *se*-stem demonstratives are used for referents that are activated and already sufficiently identified (Laury 1997; Etelämäki 2006). In our Finnish data relating to House 1, the demonstrative adverbs are used mostly for maintaining the reference while the main function of the utterance is related to something other than identifying the referent. This may be, for instance, describing details of the referent as in (11). Here, the same building (House 1) has been mentioned in several preceding utterances, as the speaker has

¹³Example (10) epitomises the difficulties of distinguishing between the pronominal and adverbial uses of the internal case forms. Following the criteria given in §3.3, the demonstrative *tässä* is unambiguously a pronoun, since it agrees with the head of the phrase in case and number. However, one could argue that *tässä* is external to the phrase and just happens to occur in the same case with *lähemmässä talossa*. The pause between *tässä* and *lähemmässä talossa* could point to such an analysis (see Laury 1996: 86 for a similar discussion). Nevertheless, (10) and others like it are not included in the present analysis.

¹⁴For comparison: the external case forms are used to refer to the buildings a total of eight times in our data (and none of these uses expresses location), whereas the internal case forms are used 335 times.

¹⁵*Se*-stem has two other variants (*sie-* and *sii-*) but for clarity they are also referred to as *se*-stem.

been describing its roof. The speaker is using the demonstrative adverb *siellä* ‘(in) there’, which has the most common anaphoric stem.

(11) (F07.011)¹⁶

- ja **siellä** on kaunis kolmiomainen
 and there.DEM.ADV.LOC be.PRS.3SG beautiful.NOM triangular.NOM
 kartiomainen (.) tiilikatto? joka on minusta aina
 conical.NOM brick.roof.NOM that be.PRS.3SG 1SG.ELA always
 ollut (.) kaunein kattoma(ili).
 be.PST.PTCP beautiful.SUPL.NOM roof.type.NOM
 ‘And there is a beautiful triangular (or) conical brick roof **there**, which I
 have always thought to be the most beautiful roof ty(pe).’

Next, we will examine the use of internal local cases when referring to House 1, see Table 7 and (9).

Table 7: Use of the internal local cases of demonstrative pronouns in referring to House 1 in Situation 1

DemPron stem	1 st mention	2 nd mention	3 rd mention and further	Total
tä-	23 (76.7%)	4 (13.3%)	3 (10%)	30 (100%)
t(u)o-	5 (71.4%)	2 (28.6%)	–	7 (100%)
se-	2 (4.4%)	12 (26.7%)	31 (68.9%)	45 (100%)

Compared to the data on the demonstrative adverbs above, the data on the use of internal local cases reveals somewhat different referential practices. Internal local cases are used for first mentions (in the referential chain) more often than adverbs. Since there are more first mentions, there are also more *tä-* and *t(u)o*-stems, as these stems are used for identifying a referent and they typically indicate that the process of identifying the referent is unfinished (Laury 1997; Etelämäki 2005). Similar to adverbs, the *se*-stem is the most frequently used stem also among the pronouns.

Demonstrative adverbs can co-occur with demonstrative pronouns in compound forms. Typically, both demonstratives in a compound form have the same

¹⁶The preceding context for the example is the following: *jos mun pitäisi kuvalla sitä ihmiseille joka ei näe sitä niin se on sanoisin että (.) se on ensin hyvin suorakulmainen, melkein neliö.* ‘If I needed to describe it to a person who cannot see it, then it is, I would say that at first it is very rectangular, almost like a square.’

stem, but compounds containing demonstratives with different stems are also possible. The frequency of compound forms is low in the Finnish data (a total of 16 referential units in the data, including all referents and both parts of the experiment). Therefore, in this part of the analysis, we will use examples concerning all three referents (Houses 1, 2 and 3). The most typical structure of a demonstrative compound is either pronoun+adverb (12) or pronoun+noun+adverb (13). It is also possible for the adverb to precede the pronoun.

- (12) (F21.031)

ja (.) draakkonin rakennus on pienempi ku
and draakkon.GEN building.NOM be.PRS.3SG small.COMP.NOM than
toi (.) toi tuolla.
that.DEM.PRON.NOM that.DEM.PRON.NOM over.there.DEM.ADV.LOC
'And the Draakkoni building is smaller than that one over there.'

- (13) (F19.001)

no: tää on luultavasti vanhempi
well this.DEM.PRON.NOM be.PRS.3SG probably old.COMP.NOM
tää (.) draakkoni tässä,
this.DEM.PRON.NOM draakkoni.NOM this.DEM.PRON.INE
'Well this one is probably older, this Draakkoni (= name of the building)
here.'

When internal local cases of pronouns appear in compound forms, their adverb-like features are particularly apparent. Unlike other pronouns, internal local cases of demonstratives can appear after the head of the noun phrase, as in (13) above, where the demonstrative pronoun *tässä* appears in a position which is typical for adverbs in compound forms.

The co-occurrence of two demonstratives reinforces the pragmatic function of the referential unit (Diessel 2006: 474). The speaker's reason for using a compound form may have to do with hesitation and uncertainty about whether the intended referent has been sufficiently identified in order to be distinguished from competing referents. The pauses within the referential units in (12) and (13) could also point in this direction, as may the dislocated position of the referential unit containing the compound form in (13) (on the role of right-dislocations in reducing disambiguation and clarifying the referent, see e.g. Geluykens 1987).

Both the adverbs and the internal case forms can be used for identifying a referent and tracking a referent. By choosing either the adverbial form or the

pronominal form of a demonstrative, the speaker can emphasise certain features of the referent. In (14), the speaker first refers to House 2 (the more distant building) with a distal adverb *tuol* ('over there'). Then follows a reference to House 1 (the closer building) by using the same distal stem but this time in a pronominal form (*tos* 'in that'). At this point, the contrast between the two referents is expressed by using an adverb for one and a pronoun for the other referent. At the end of the utterance, the speaker refers to House 1 once more with an expression that contains an NP with a demonstrative pronoun, this time with a proximal stem (*tässä* 'in this').

- (14) (F21.015)

(.)	noi	ei	oo	savupiiput	ku
	that.DEM.PRON.NOM.PL	NEG.3PL	be.PRS	chimney.NOM.PL	that
tuol=	on	mut	tos=	on	
over.there.DEM.ADV.LOC	be.PRS.3SG	but	that.DEM.PRON.INE	be.PRS.3SG	
vanha	savupiippu	tässä		draakkonis.	
old.NOM	chimney.NOM	this.DEM.PRON.INE	draakkoni.INE		

'Those **over there** are not chimneys, but **that one** has an old chimney,
this Draakoni.'

Keeping in mind the list of possible factors affecting the choice of the referential device (exactness of reference, visibility of the referent, proximity of the referent and size or boundedness of the area referred to; see §3.3) we can try to explain the referential strategies the speaker is employing here. In this context, House 1 and its chimney are more proximal, more visible and more exact than House 2 and the elements on its roof (that the speaker apparently cannot fully identify). This causes the speaker to choose a pronominal form for House 1 and an adverb for House 2.¹⁷ To emphasise the relative proximity of House 1 and to ensure that the switch between referents will be correctly interpreted, the speaker produces a dislocated referential unit that includes a proximal pronoun. This unit also includes a lexical NP in order to identify the intended referent even more clearly.

¹⁷The choice of the demonstrative form also creates a syntactic difference between the sentences about House 1 and House 2. The use of the pronominal form in the second sentence (... *tos on vanha savupiippu* ...) makes it possible to interpret it as describing a part–whole relationship. (The referential unit refers to the building in order to describe an entity that is part of the building.) This kind of interpretation would not be possible with the first sentence, where the adverbial form serves only to locate the elements the speaker wishes to describe. We thank the anonymous reviewer for pointing this out.

It is important to note that the choice of demonstrative form is not related to the objective features of the referent alone. Instead, it has more to do with how the referent is conceptualised in that particular context (Laury 1996: 83–84). This can be observed in (15), where the speaker refers to House 1 first with a pronominal form *siinä* ‘in it/that’ and after that twice with an adverb *siel/sielä* ‘there’. The first referential unit points to the entire building, whereas the two subsequent referential units refer to parts of the building instead.

- (15) (F19.025)

siinä on hauskasti muutten laitettu
it.DEM.PRON.INE be.PRS.3SG funny.ADV by.the.way set.PST.PTCP
noi (.) sadevesijärjestelmät et
that.DEM.PRON.NOM.PL rain.water.system.NOM.PL that
siel= on (.) tosiaan draakkonin hampaat
there.DEM.ADV.LOC be.PRS.3SG indeed draakkon.GEN tooth.NOM.PL
sielä (.) molemmissa.
there.DEM.ADV.LOC both.INE

‘In that one, they have installed the rainwater systems in a funny way, so that there are indeed teeth of a *draakkon* (here: a dragon) there in both of them.’

At first, the speaker is using a pronominal form to confirm that the referent is still the same as in the preceding two utterances (also related to House 1). When the speaker starts to describe the details of the referent and the focus is no longer on identifying the referent, he switches to a different referential device. The adverbs serve to locate the part of the referent that is central for the description.

To summarise the differences in usage between the demonstrative adverbs and the internal case forms of the demonstrative pronouns in the Finnish data, the following tendencies can be detected: the demonstrative adverbs are mainly used i) for tracking a referent, ii) for expressing location alone, iii) in contexts where the focus of current activity is somewhere else and not focused on identifying a referent. Pronominal forms, on the other hand, are used i) for identifying a referent, ii) for both locative and (at least partially) non-locative concepts, such as expressing a part–whole relationship, iii) in contexts where the process of identifying the referent is unfinished and the referential act requires precision. The demand for precision is related to multiple features of the situation: the physical features of the referent on one hand and discourse-level features on the other.

The internal local cases of pronouns are used more frequently than the adverbs in our Finnish data. The internal local cases carry a comprehensive set

of functions that relate to both the spatial and discourse features of the referent. This makes them an optimal referential device for a spatially contrastive context where the speaker needs to concentrate on keeping two or more competing referents apart, while simultaneously giving detailed information on the referents. That being said, the pragmatic functions carried by the pronominal forms on one hand and the adverbial forms on the other, have developed from the tension between two parallel and partially overlapping systems for spatial reference: a pronominal system and an adverbial system. To attain maximum referential power, speakers use both systems in parallel and simultaneously in different combinations.

5 Conclusions

The main aim of this chapter was to explore the functions of demonstrative adverbs in a spatially contrastive setting and the relationship between the use of demonstrative adverbs in our data and the general demonstrative system (including pronouns) that the languages in our sample employ. The analysis of the experimentally elicited spoken data from Russian, Estonian and Finnish resulted in the following conclusions:

i) In all three languages, the proximal demonstrative adverbs occur mostly in first mentions of the referential chain, whereas the distal ones (or, in Finnish, mostly hearer-centred ones) occur when the referent is activated (that is, the referent has already been mentioned in the preceding utterance[s]). The proximal demonstratives can be used anaphorically as well, but this kind of usage is not frequent in our data. This suggests that similar to what Gundel et al. (1993; 2010) have proposed for demonstrative pronouns, demonstrative adverbs can indicate different activation statuses in the givenness hierarchy. To conclude, in our Estonian and Russian data, distal demonstrative adverbs serve an anaphoric function indicating that the referent is activated/in focus. In our Finnish data, this function is most commonly conveyed through using a hearer-centred demonstrative.

ii) In Finnish, demonstrative adverbs are used for tracking referents that are already activated. Their main function is to maintain the reference while the speaker is focusing on some other activity – for instance, describing the referent in detail. In the Finnish data, the internal local cases of demonstrative pronouns occur in first mentions of the referential chain more often than adverbs. These pronouns are used for identifying a referent and in contexts where the referential act requires exactness or when the process of identifying the referent is unfinished. These observations are in line with Laury's (1996) view according to

which referents referred to with the internal local cases tend to be conceptualised as figures and those referred to with adverbs to be conceptualised as ground.

As in the Finnish data, demonstrative adverbs are used in our Russian data mainly for tracking already activated referents. To identify the referent, Russian speakers rely on bare noun phrases more often than Finnish speakers.

Furthermore, in Finnish and Estonian, demonstrative pronouns and demonstrative adverbs co-occur in compound forms. Similar to the internal local cases of pronouns in Finnish, the compound forms (that is, the forms containing both a demonstrative pronoun and a demonstrative adverb) are used to indicate that the referent is not yet sufficiently defined. Demonstrative adverbs do not co-occur with demonstrative pronouns in the Russian data.

The frequency of different referential devices in the data might have a connection to the general demonstrative systems employed by the languages in our sample. Finnish, having a demonstrative system with the largest number of spatial distinctions, displays the greatest proportion of demonstratives (pronouns included) in the data. Moreover, the Finnish demonstratives exhibit the most fine-tuned division of labour between pronouns and adverbs. Russian has the smallest proportion of demonstratives in the sample, and the use of demonstratives is syntactically more limited in Russian than in Finnish or Estonian (Russian demonstrative adverbs do not function as determiners or occur in compound forms with demonstrative pronouns). Instead, the Russian data contains the largest proportion of personal pronouns and the second largest proportion of noun phrases without a demonstrative. Estonian is located “between” Finnish and Russian in many ways. The Estonian demonstrative system is less elaborate than the Finnish one; nevertheless, the Estonian demonstratives are syntactically more flexible than the Russian ones, as they can also occur in compound forms with pronouns. The Estonian data contains the greatest proportion of demonstrative adverbs but also the greatest proportion of bare noun phrases.

The present analysis suggests that the choices a speaker makes in reference resolution in a spatially contrastive setting may be related to properties of the demonstrative system, as well as to the overall system of the language’s referential devices and demonstrative practice. However, the present study is based on a small sample of relatively homogeneous data, which does not cover all aspects of demonstrative use. Therefore, we can only make tentative observations on the correlation between size of a deictic system and frequency of demonstratives in a spatially contrastive setting. Further work is needed to examine more closely the links between properties of the demonstrative system and demonstrative practice.

Acknowledgments

Research for this chapter has been supported by the Estonian Research Council grant PUT 701, and by the (European Union) European Regional Development Fund (Centre of Excellence in Estonian Studies).

Abbreviations

Symbols used in the transcription (Jefferson 2004)

=	end of one unit and beginning of the next one without a pause in between	.	falling intonation
(.)	brief pause	,	flat intonation
:	prolonged vowel or consonant	?	rising intonation
		(word)	an uncertain unit or part of a unit

Abbreviations used in the glosses

1	first person	NEG	negation
3	third person	NOM	nominative case
ACC	accusative case	PART	partitive case
ADE	adessive case	PERF	perfect tense
ADV	adverb	PL	plural
ALL	allative case	PREP	prepositional case
COMP	comparative	PRON	pronoun
DEM	demonstrative	PRS	present tense
DIM	diminutive	PRTCL	particle
ELA	elative case	PST	past tense
GEN	genitive case	PTCP	participle
ILL	illative case	REL	relative pronoun
INE	inessive case	SG	singular
INF	infinitive	SUPL	superlative
LOC	locative (case)		

References

- Anderson, Stephen R. & Edward L. Keenan. 1985. Deixis. In Timothy Shopen (ed.), *Language typology and grammatical description*, vol. 3: *Grammatical categories and the lexicon*, 259–308. Cambridge: Cambridge University Press.
- Ariel, Mira. 2013. Centering, accessibility and the next mention. *Theoretical Linguistics* 39(1–2). 39–58.
- Coventry, Kenny R., Berenice Valdés, Alejandro Castillo & Pedro Guijarro-Fuentes. 2008. Language within your reach: Near–far perceptual space and spatial demonstratives. *Cognition* 108(3). 889–895.
- Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Diessel, Holger. 2013. Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. *Language and Cognition* 5(2–3). 239–249.
- Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1). 61–112.
- Duvallon, Outi. 2005. The pronoun *se* in the context of syntactic and discursive ruptures of spoken texts. In Ritva Lauri (ed.), *Minimal reference: The use of pronouns in Finnish and Estonian discourse*, 163–189. Helsinki: Finnish Literature Society.
- Enfield, Nick. 2003. Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis. *Language* 79(1). 82–117.
- Etelämäki, Marja. 2005. Context and referent in interaction. In Ritva Laury (ed.), *Minimal reference: The use of pronouns in Finnish and Estonian discourse*, 12–37. Helsinki: Finnish Literature Society.
- Etelämäki, Marja. 2006. *Toiminta ja tarkoite: Tutkimus suomen pronominista tämä* [The use and meaning of the Finnish pronoun *tämä*]. Helsinki: University of Helsinki.
- Geluykens, Roland. 1987. Tails (right-dislocations) as a repair mechanism in English conversation. In Jan Nuyts & Georges de Schutter (eds.), *On word order and functional grammar*, 119–129. Dordrecht: Foris.
- Grenoble, Lenore A. 1998. *Deixis and information packaging in Russian discourse*. Amsterdam: Benjamins.
- Gudde, Harmen B., Kenny R. Coventry & Paul E. Engelhardt. 2016. Language and memory for object location. *Cognition* 153. 99–107.

- Gundel, Jeanette K., Mamadou Bassene, Bryan Gordon, Linda Humnick & Amel Khalfaoui. 2010. Testing predictions of the givenness hierarchy framework: A crosslinguistic investigation. *Journal of Pragmatics* 42(7). 1770–1785.
- Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307.
- Hakulinen, Auli. 1985. On cohesive devices in Finnish. In Emil Sözer (ed.), *Text connexity, text coherence: Aspects, methods, results*, 337–362. Hamburg: Buske.
- Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen & Irja Alho. 2004. *Iso suomen kielioppi* [The large grammar of Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Halliday, M. A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Routledge.
- Hanks, William F. 1990. *Referential practice: Language and lived space among the Maya*. Chicago, IL: University of Chicago Press.
- Hanks, William F. 1992. The indexical ground of deictic reference. In Alessandro Duranti & Charles Goodwin (eds.), *Rethinking context: Language as an interactive phenomenon*, 43–76. Cambridge: Cambridge University Press.
- Hanks, William F. 2011. Deixis and indexicality. In Wolfram Bublitz & Neal R. Norrick (eds.), *Foundations of pragmatics*, 315–346. Berlin: De Gruyter Mouton.
- Himmelmann, Nikolaus. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora*, 205–254. Amsterdam: Benjamins.
- Itkonen, Terho. 1966. Tutkimus suomen asyndetonista [Research on asyndeton in Finnish]. *Virittäjä* 70(4). 402–423.
- Jefferson, Gail. 2004. Glossary of transcript symbols with an introduction. In Gene H. Lerner (ed.), *Conversation analysis: Studies from the first generation*, 13–31. Amsterdam: Benjamins.
- Kibrik, Andrej A. 2011. *Reference in discourse*. Oxford: Oxford University Press.
- Larjavaara, Matti. 1985. Suomen demonstratiivisysteemin rakenne [the structure of the Finnish demonstrative system]. *Sananjalka* 27. 15–31.
- Larjavaara, Matti. 1990. *Suomen deiksis* [Finnish deixis]. Helsinki: Finnish Literature Society.
- Laury, Ritva. 1996. Pronouns and adverbs, figure and ground: The local case forms and locative forms of the Finnish demonstratives in spoken discourse. *SKY Yearbook of the Linguistic Association of Finland* 9. 65–92.
- Laury, Ritva. 1997. *Demonstratives in interaction: The emergence of a definite article in Finnish*. Amsterdam: Benjamins.
- Levinson, Stephen C. 2004. Deixis. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 97–121. Malden, MA: Blackwell.

- Levinson, Stephen C. 2018. Introduction: Demonstratives: Patterns in diversity. In Stephen Levinson, Sarah Cutfield, Michael J. Dunn, N. J. Enfield & Sérgio Meira (eds.), *Demonstratives in cross-linguistic perspective*, 1–42. Cambridge: Cambridge University Press.
- Levinson, Stephen C., Sarah Cutfield, Michael J. Dunn, N. J. Enfield & Sérgio Meira (eds.). 2018. *Demonstratives in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Maes, Alfons & Christ Rooij. 2007. (How) do demonstratives code distance? In *Proceedings of the 6th Discourse Anaphora and Anaphora Resolution Colloquium DAARC 2007*, 83–89. Lagos: Centro Linguistica da Universidade do Porto.
- March, Evrim G. & Philippa Pattison. 2014. The role of language system in context-dependent language use in Turkish-speaking versus English-speaking older adults. *Applied Psycholinguistics* 35(6). 1087–1108.
- Meira, Sérgio & Angela Terrill. 2005. Contrasting contrastive demonstratives in Tiriýó and Lavukaleve. *Linguistics* 43(6). 1131–1152.
- Östman, Jan-Ola. 1995. Recasting the deictic foundation, using physics and Finnish. In Masayoshi Shibatani & Sandra Thompson (eds.), *Essays in semantics and pragmatics: In honor of Charles J. Fillmore*, 247–278. Amsterdam: Benjamins.
- Pajusalu, Renate. 2006. Death of a demonstrative: Person and time. The case of Estonian *too*. *Linguistica Uralica* 42(4). 241–253.
- Pajusalu, Renate. 2009. Pronouns and reference in Estonian. *STUF – Language Typology and Universals / Sprachtypologie und Universalienforschung* 62(1–2). 122–139.
- Pajusalu, Renate. 2017. Viiteseosed. In Mati Erelt & Helle Metslang (eds.), *Eesti keele süntaks* [The syntax of Estonian], 566–589. Tartu: Tartu Ülikooli Kirjastus.
- Reile, Maria. 2015. Space and demonstratives: An experiment with Estonian exophoric demonstratives. *Eesti ja soome-ugri keeleteaduse ajakiri / Journal of Estonian and Finno-Ugric Linguistics* 6(2). 137–165.
- Reile, Maria. 2016. Distance, visual salience, and contrast expressed through different demonstrative systems: An experimental study in Estonian. *SKY Journal of Linguistics* 29. 63–94.
- Reile, Maria, Piia Taremaa, Tiina Nahkola & Renate Pajusalu. 2019. Reference in the borderline of space and discourse: Free production experiment in Estonian, Finnish and Russian. *Linguistica Uralica* 55(3). 185–208.
- Sahkai, Heete. 2003. Demonstrative doubling in Spoken Estonian. *Trames* 7(2). 120–144.

- Seppänen, Eeva-Leena. 1998. *Läsnäolon pronominit - tämä, tuo, se ja hän viittaa-massa keskustelun osallistujaan* [Pronouns referring to participants in the conversation]. Helsinki: Finnish Literature Society.
- Sheljakin, M. A. 2002. *Russkij jazyk: Spravochnik* [Russian language: A handbook]. Tallinn: Koolibri.
- Siiotonen, Kirsti. 1979. Lokaalisten pronominaaliadverbien merkityseroista suomessa ja saksassa [Regarding differences in the meaning of local pronominal adverbs in Finnish and German]. In *Sanomia. A festschrift in honor of the 60th birthday of Eeva Kangasmaa-Minn*. Turku: University of Turku.
- Silverstein, Michael. 1976. Shifters, linguistic categories and cultural description. In Keith H. Basso & Henry A. Selby (eds.), *Meaning in anthropology*. Albuquerque: University of New Mexico Press.
- Slobin, Dan I. 1996. From "thought and language" to "thinking for speaking". In John J. Gumperz & Stephen C. Levinson (eds.), *Rethinking linguistic relativity*, 70–96. Cambridge: Cambridge University Press.
- Talmy, Leonard. 1983. How language structures space. In Herbert Pick & Linda Acredolo (eds.), *Spatial orientation: Theory, research and application*. New York, NY: Plenum.
- Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Chapter 11

Manner deictics in quotative indexes of Finno-Ugric

Denys Teptiuk

University of Tartu

This chapter discusses the quotative use of manner deictics in computer-mediated communications of five languages representing three Finno-Ugric branches: Finnic, Permic and Hungarian. The aim of the study is to (i) define the functional properties of manner deictics in quotative indexes (QIs) of the languages in focus, (ii) demonstrate in what types of QIs they appear, and (iii) determine possible functional similarities in the distribution of the markers between the languages. It is shown that manner deictics can be employed as cataphoric (Finnish *näin*, *niin*, *sillee(n)*; Estonian *nii*; Hungarian *úgy*; Udmurt *tažy*) and anaphoric (Udmurt *ožy*) markers, or can function as both (Komi *tadž(i)*, *sidž*; Hungarian *így*). Furthermore, some manner deictics (Finnish *sillee(n)*, Hungarian *így*) introduce mimetic expressions that can be interpreted as quasi-quotations. In the conclusion, cross-linguistic parallels in the use of manner deictics in quotative constructions are pointed out.

1 Introduction

Recent cross-linguistic studies have shown that comparative/similative *like*, demonstrative deictic *so* and quantifying elements *just* and *all*, as well as motion (*go*) and action (*do*) verbs, can grammaticalise into quotative markers (see Buchstaller & van Alphen 2012: xii–xiv; Güldemann 2008: §5.1.2–§5.1.5). Quotative markers with demonstrative semantics are found in many typologically diverse languages. Manner deictics are either used in quotative constructions together with reportative verbs, i.e. speech or epistemic verbs, or they are used alone to point out the presence of a quote (Güldemann 2008: 321, 350). Their use is explained by the function of quotations as a type of demonstration embedded in language use, i.e. “a mimetic reenactment of a non-immediate state of affairs”



(Güldemann 2008: 320). By producing an utterance from a different temporal or spatial situation, the reporter demonstrates the situation to an audience (Clark & Gerrig 1990: 802; see Clark 2016 on speech acts as demonstrations). The endophoric use of demonstratives in quotative constructions is usually associated with a cataphoric reference, which “relate[s] to stretches of following discourse” (König & Umbach 2018: 297; see also Güldemann 2008). Although cataphoric reference seems to be the most commonly attested extended use of manner deictics in the world’s languages, this statement might be too general. Consider, for example, (1) from Hungarian, in which the manner deictic *így* ‘so’ is used anaphorically, pointing at the just-produced quote.

- (1) Hungarian (MNSz)¹
“Nagy pénz, kis foci.”_{RD} – vagy mégsem így mondta
big money small football or still.also.NEG so say.PST.3SG.DEF
volna Puskás?
be.COND.3SG PN
‘“Big money, small football.”_{RD} – as Puskás would have said (lit. or
wouldn’t have Puskás said so?)’²

Hence, there is a reason to look more closely at the direction of endophoric reference of manner deictics in quotative constructions. For this purpose, I conducted a contrastive study on the quotative use of manner deictics in computer-mediated communications of five distantly related Finno-Ugric languages. The choice of languages is not accidental. They belong to three different branches of the language family, i.e. Finnic (represented by Finnish and Estonian), Permic (Komi and Udmurt), and Hungarian, and to three different geographical areas: Northern Europe (Finnish and Estonian), Central Europe (Hungarian) and Russia (Komi and Udmurt). Despite their relatedness and typological closeness, the languages did not have contact with each other for centuries, with the exception of Finnish and Estonian, and they possess individual features that developed independently or through contact with languages in their respective areas. Hence, I suspect these languages represent different typological patterns in the use of manner deictics in quotative constructions which can be determined and further applied cross-linguistically. By taking a closer look at the distribution of

¹See §3 for details on the type of data used in the study. The abbreviated sources in parentheses in the first line of every example are references to the list of electronic resources given in the appendix.

²In the examples, quotative indexes are marked in bold, and reported discourse is enclosed in double vertical bars.

manner deictics in the five Finno-Ugric languages, I aim to determine their functional properties in quotative constructions, e.g. do manner deictics introduce only particular types of reported speech and thought or can they function as general mimetic markers? Do structural features have an impact on their functions? Is there any correspondence between their use inside and outside of quotative constructions? Although descriptive grammars (e.g. Erelt & Metslang 2017 on Estonian; Hakulinen et al. 2004 on Finnish) and previous studies provide basic descriptions of manner deictics in individual languages and even touch upon their quotative use (e.g. Keevallik 2005 on Estonian; Kiefer 2016 on Hungarian), these and other questions in relation to their quotative use still remain unexplored.

The chapter is organised as follows. In §2, I provide the terminological framework for this study. In §3, my methodology and database are described. §4 presents the results of the study of Finnish and Estonian (§4.1), Komi and Udmurt (§4.2), and Hungarian (§4.3). Finally, §5 summarises the main findings and highlights cross-linguistic similarities in the quotative use of manner deictics based on the determined typological patterns.

2 Terminological framework

In my investigation on manner deictics in five Finno-Ugric languages, I adapt Güldemann's framework of reported discourse (henceforth RD), which is defined as follows:

Reported discourse is the representation of a spoken or mental text from which the reporter distances him-/herself by indicating that it is produced by a source of consciousness in a pragmatic and deictic setting that is different from that of the immediate discourse (Güldemann 2008: 6).

Güldemann prefers “discourse” as “the representation of spoken or mental text” over the more traditional “speech”, since RD “is not restricted to real instances of speech” and may also include “texts that were never actually uttered like so-called ‘internal speech’, or in general any representation of cognitive acts or states” (Güldemann 2008: 7).

According to Güldemann (2008: 10), RD together with the elements introducing it form a complex whole labelled as an “RD-construction”. RD-constructions canonically consist of two major constituents: RD and “quotative index” (henceforth QI). In (2), the clause *I said to him* is a QI followed by the RD “*Your party won't be pleased ...*”.

- (2) English (Daily Mail)

I said to him ||“Your party won’t be pleased ...”||_{RD}

Güldemann (2008: 11) defines a QI as “a segmentally discrete linguistic expression which is used by the reporter for the orientation of the audience to signal in his/her discourse the occurrence of an adjacent representation of reported discourse”. As a “linguistic expression”, QIs can represent structures of different complexities, consisting of “just a gram (...), an independent function word, a phrase, a full sentential syntagm (...), or even a clause with more than one predicate” (Güldemann 2008: 11).³ Example (2) contains only one of various possible QI forms, namely a speech verb (*said*) describing the event behind the RD (the quotation of speech), and NPs encoding participants, i.e. the original speaker (*I*) and an addressee (*to him*). However, although these constituents are probable and relatively frequent they are not indispensable elements of the QI. Even speech verbs are not universal components of QIs per se. Consider (3), for instance, where instead of a speech or epistemic verb (henceforth labelled reportative verb), the reporter uses the motion verb *go* and the combination of the equational verb *be* with the simulative marker *like* for the presentation of RD.

- (3) English (Twitter)

... and he goes ||I am the police bitch||_{RD} and starts touching the register
I’m like ||oh this mf’er didn’t just do that.||_{RD}

Manner deictics as non-reportative elements are expected to co-occur primarily with reportative elements in QIs, as in (1). However, their co-occurrence with other (grammaticalised) elements is also investigated here. Furthermore, it is also of interest to see whether they undergo changes in the quotative domain and grammaticalise into genuine quotative markers that are not bound to reportative elements. Therefore, I pay attention to the use of manner deictics in different constructions and point out additional meanings and functions that can be observed in their use in less or more complex QIs.

³In some situations, QIs can remain verbally unexpressed. Instead, suprasegmental features of intonation, dynamics and pitch can be used as the sole means to contrast a quote with its surrounding contexts. Since I use non-standard written texts as a corpus for this study (see §3) and focus on the use of manner deictics in quotative constructions, I exclude verbally unexpressed QIs from the current investigation.

3 Methodology and data

In this study, data originating from social network sites (SNS) is used as a database. My choice of SNS data is motivated by the following factors. First, the focus is on QIs that are mainly used in non-standard varieties of the studied languages, which typically exceed the limited amount of QIs used in standardised texts. Second, previous studies show that “[t]he informal characteristics of SNS enables the usage of generally oral forms such as slang and dialects in a written context” (Pischlöger 2014: 144). As for the minority languages Komi and Udmurt, “the relaxed atmosphere on SNS allows language use which is typical for oral communication and otherwise frowned upon in other (especially written) contexts by language purists” (Pischlöger 2014: 144). Thus, Komi and Udmurt speakers use language online that not only includes variants that are commonly mixed with the dominant Russian language, but also a mixture of dialects and styles that are peculiar to colloquial speech (see Pischlöger 2016; Edygarova 2013, 2014). Thus, despite the presence of emoticons, different orthographic symbols and nonstandard shortenings, the language on SNS can be considered a written approximation of spoken language, combining the features of colloquial speech and standard writing in one text (Helasvuo et al. 2014).

For data collection, I studied the occurrence of quotations in different new media sources. Since Komi and Udmurt are endangered languages with a smaller amount of online material compared to Finnish, Estonian and Hungarian, I also investigated available text collections for the Permic languages, i.e. Uotila (1985; 1989) for Komi and Kel'makov (1981; 1990) for Udmurt. The collections provide transcribed oral narratives by Komi and Udmurt speakers from various dialectal groups. This material supplemented the data from new media sources if the latter did not provide sufficient evidence. For Udmurt, the material predominantly originates from the group *Jumshan57*⁴ on the Russian SNS *vk.com*. I studied approximately 100 of the 249 pages (ca. 200 blog posts) of the material available. In addition, I used the blog page *KYLZY Jopte*⁵ (containing 35 pages and 696 blog posts at the time of investigation) as well as other groups and pages. The choice of groups was motivated by four factors: (i) the number of entries, (ii) the number of group members, (iii) the use of Udmurt (exclusive or parallel to Russian), and (iv) the dominance of unedited texts. Thus, I preferred unofficial pages consisting of live conversations in comment sections and unedited blog entries to official pages of media resources, non- and governmental organisations, etc. As a result,

⁴<https://vk.com/knyazpozdey> (last accessed August 1, 2019).

⁵https://vk.com/udmurt_ept (last accessed August 1, 2019).

I could also trace the systematic appearance of vernacular variants mixed with Russian in live discussions in comment sections, which would have been impossible with standardised texts. Where additional material was required, data from the Blog subcorpus of the Udmurt corpus (henceforth Blog subcorpus)⁶ was used to further investigate the use of manner deictics in QIs.⁷ The Blog subcorpus contains approximately 160 examples of the proximal manner deictic *tažy* and approximately 1400 examples of the distal manner deictic *ožy* (see §4.2 on the use of Udmurt manner deictics). Where the number of examples was fairly low (as in the case of *tažy*), I checked every instance separately. In the case of a higher number of examples, I checked the collocation of this marker with the reportative verbs *šuyny* ‘say’ and *malpanы* ‘think’ in different finite forms, e.g. ‘(I) say/said thus’, ‘(she/he) thinks/thought so’, etc. With the combination of the new media sources and the online corpora I collected approximately 40 examples of manner deictics in quotative constructions and 20 examples of their use outside the quotative domain.

Taking the amount and quality of Udmurt data as a reference point, I studied a similar number and type of pages on *vk.com* and *blogspot.com* for Komi. However, at the end of the investigation, my corpus contained only three examples of the manner deictic *tadž(i)* in quotative constructions. Therefore, I used the corpus of the Komi language (henceforth Komi corpus)⁸ as a supplementary source to make generalisations about the use of manner deictics in QIs. The corpus contains oral and written texts of various genres (fiction, journalistic texts, educational and scientific literature, official correspondence, etc.), excluding new media texts.

For Finnish, I used the Corpus of Internet Communications,⁹ consisting of data from the forums *Suomi24* and *Ylilauta*,¹⁰ as my main material. I browsed the corpus for collocations of the manner deictics *näin*, *niin*, *noin*, *tälläin* (~ *tälläi*), *t(u)olloin* (~ *tollai*) and *sillee(n)* (~ *sillai*) with the reportative verbs *sanoa* ‘say’ and *ajatella* ‘think’, and the equational verb *olla* ‘be’. Based on my previous investigations (Teptiuk 2019), these verbs were expected to be the most probable

⁶<http://web-corpora.net/UdmurtCorpus> (last accessed August 1, 2019). The Blog subcorpus consists of ca. 6% of the whole corpus (7.3 million tokens).

⁷At the time of data collection, the Volga-Kama Udmurt corpora containing the new media subcorpus (<http://volgakama.web-corpora.net>, last accessed August 1, 2019) were not available.

⁸The corpus of the Komi language (Russian: *Korpus komi jazyka*) is available at: <http://komicorpora.ru> (last accessed July 1, 2019). An exact number of tokens in the corpus is not specified.

⁹Finnish: *Internet-keskusteluaineistoja*.

¹⁰<https://korp.csc.fi> (last accessed July 1, 2019). The corpus consists of ca. 6.9 billion tokens.

components of Finnish QIs. I checked the first 100 examples of every query for the quotative use of these collocations when the number of queries was too high.

In addition, I used the Google search engine, which allowed me to test these collocations on web pages outside the digital corpus when the number of collocations was low. I checked the first ten pages of the Google search results, with each page containing ten links. Since the search results are less accurate the further one proceeds, I considered this amount to be enough to investigate the use of manner deictics in QIs (ca. 100 search results for one tested variant of a construction). In order to make the search more effective, I placed the studied material into quotation marks and checked the different collocations of reportative verbs and manner deictics in their different grammatical forms. Pages that did not fall under the category of computer-mediated communications (CMC), e.g. edited newspaper articles or science fiction texts (see Crystal 2001), were not taken into account as primary sources of examples. I used data outside CMC only when primarily suitable sources did not yield any relevant results. In all, I collected approximately 50 examples of Finnish manner deictics in quotative constructions.

For Estonian and Hungarian, I used both suitable corpus data (Estonian: New media subcorpus of the Estonian Reference Corpus,¹¹ henceforth NMS; Hungarian: Personal subcorpus of the Hungarian National Corpus,¹² henceforth MNSz) and the independently collected material obtained through Google searches. I checked for the collocations of manner deictics with reportative verbs within the corpora and Google searches. For Estonian, collocations with the equational verb *olema* ‘be’ were also considered (see Teptiuk 2019 for more details on Estonian QIs). Altogether, I collected approximately 20 examples for Estonian and approximately 30 examples for Hungarian.

All examples are provided here with translations and glosses. Minor spelling mistakes are corrected, but punctuation errors are not. Since Komi and Udmurt use different symbols to mark identical sounds, I present these in the transcriptions to avoid confusion while transliterating the Permic examples. Russian code-switches are presented in transliteration and enclosed in curly brackets {} in the glosses.

¹¹<http://www.cl.ut.ee/korpused/segakorpus/> (last accessed July 1, 2019); the subcorpus contains ca. 21 million words.

¹²<http://corpus.nytud.hu/mnsz> (last accessed July 1, 2019); the subcorpus contains 18.6 million words of discussion on internet forums, mainly deriving from the Hungarian internet portal, *index.hu*, and several forums from *Subcarpathia* (RO, UA).

4 Manner deictics as quotatives in Finno-Ugric languages

4.1 Manner deictics in quotative indexes in Finnish and Estonian

Finnish has a tripartite system of manner deictics: (i) speaker-proximal *näin*, (ii) hearer-proximal *noin*, and (iii) distal *niin* (König 2017: 147; also see Hakulinen et al. 2004: §668). The manner deictics historically derive from the instructive case forms of the plural demonstratives *nämä*, *nuo* and *ne*. Besides their basic functions as manner deictics, they are used as intensifiers (Hakulinen et al. 2004: §792), or, in the case of *niin*, as a correlative pronoun (Hakulinen et al. 2004: §1160, §1172).

In addition, one can find another set of manner deictics used in colloquial Finnish (Hakulinen et al. 2004: §721) with the same distinction in the deictic domain: *tällee(n)*, *t(u)ollee(n)* and *sillee(n)*. These forms historically derive from the singular demonstratives *tämä*, *tuo* and *se* in the allative case (-*lle*) and the 3rd singular possessive suffix (-*Vn*). *Tällee(n)* and *t(u)ollee(n)* are rarely mentioned in previous descriptions and probably are used less often in contemporary Finnish. *Sillee(n)* has the same basic demonstrative functions as *niin* (Hakulinen et al. 2004: §1160). Additionally, it is used as a discourse particle in self-repairs and during speech planning (Hakulinen et al. 2004: §861). Table 1 summarises manner deictics in Finnish.

Table 1: Manner deictics in Finnish

Deictic differentiation	Standard	Colloquial
Speaker-proximal	<i>näin</i>	<i>tällee(n) ~ tallai</i>
Hearer-proximal	<i>noin</i>	<i>t(u)ollee(n) ~ tollai</i>
Distal	<i>niin</i>	<i>sillee(n) ~ sillai</i>

In quotative constructions, only *näin*, *niin* and *sillee(n)* appear systematically. Hearer-proximal *noin* is not observed in the quotative domain in 100 randomly selected examples testing its collocation with the reportative verbs *sanoa* ‘say’ and *ajatella* ‘think’ or the equational verb *olla* ‘be’ (see §3). As for the manner deictics *tällee(n)* or *t(u)ollee(n)*, their quotative use is marginal, accounting for only a couple of occurrences in the Finnish new media corpus. Therefore, I exclude them from further discussion and concentrate on *näin*, *niin* and *sillee(n)*.

In standard Estonian, in turn, only *nii*, the cognate of Finnish distal *niin*, is used regularly in quotative constructions. Outside the quotative domain, the manner

deictic *nii* also serves as an intensifier, a causal conjunction and a correlative word (Keevallik 2005: 109). The deictic distinctions of manner and other demonstratives are neutralised at the level of the literary standard. Other manner deictics, e.g. the synonymous adverbs *nõnda*, *sedamoodi*, *selliselt*, meaning ‘this way, so’, appear infrequently as synonyms to *nii* in quotative constructions. The contrastive manner deictic *naa* ‘that (other) way’ does not appear in QIs at all. Therefore, only *nii* is included in the discussion.

Finnish and Estonian manner deictics are used with cataphoric reference. They are part of the preposed QI, which includes reportative verbs. The manner deictic points at a following stretch of RD, as in (4) and (5). In constructions with reportative verbs, manner deictics are not necessary parts of the QI, and even if one omits them from the QI-clause, the RD is still understood as such. Hence, the main quotative function is carried out by reportative verbs, while the manner deictic is secondary.

(4) Finnish

- a. (Internet-keskusteluaineistoja)

... **sanoi** **näin** ||“tykkäään susta paljon, mutta en
say.PST.3SG so¹³ like.PRS.1SG 2SG.ELA a.lot but NEG.1SG
sillai.”||RD
thus

‘... he said (lit. said so) ||“I like you a lot, but not that way.”||RD’
- b. (Internet-keskusteluaineistoja)

Hän ajatteli **niin, että** ||aika on neljäs
3SG think.PST.3SG thus COMP time be.PRS.3SG four.ORD
ulottuvuuks ...||RD
dimension

‘He thought (lit. thought thus) that ||time is the fourth dimension ...||RD’
- c. (Demi)

... **sanoin** **sillee et** ||sä tykkääät must ...||RD
say.PST.1SG thus COMP 2SG like.PRS.2SG 1SG.ELA

‘... I said (lit. said thus that) ||you like me ...||RD’

¹³For the sake of convenience, here and in the glosses of other examples, I translate the proximal manner deictics as ‘so’ and the distal as ‘thus’, even though this does not reflect the actual usage of these terms in English.

(5) Estonian (Eestimaa Loomakaitse Liit)

Kaija ise ütles nii: ||“Sellist hirmu pole enam ammu
 PN self say.PST.3SG so such.PTV fear.PTV NEG more long.ago
 ühegi kassi silmis näinud ...||_{RD}
 one.GEN.PTCL cat.GEN eye.PL.INE see.PP

‘Kaija herself said (lit. said so): ||“I haven’t seen such fear in the eyes of a single cat for a long time ...||_{RD}’

Besides reportative verbs, Finnish *sillee(n)* and Estonian *nii* are also used with the equational verbs *olla* ‘be’ (6) and *olema* ‘be’ (7), respectively. Finnish *näin* appears in such constructions in only a few instances, and *niin* not at all. Since the use of equational verbs in QIs is a typical strategy in colloquial speech, the more colloquial *sillee(n)* is preferred over other manner deictics in such a construction. The same seems to be the case in Estonian. Whereas one can find examples in which *nõnda* co-occurs with a reportative verb, combinations with *olema* ‘be’ are unattested.

(6) Finnish

a. (lansiuusimaa.fi)

... se oli sillee et ||ou nou||_{RD}, ja minä olin sillee
 DEM be.PST.3SG thus COMP oh no and 1SG be.PST.1SG thus
 tyyliin, ||onks pakko||_{RD} ...
 style.ILL/like be.PRS.3SG.QP obligatory

‘... (s)he was like (lit. **was thus that**) ||oh, no||_{RD}, and I was like (lit. **was thus like**) ||is it obligatory||_{RD} ...’

b. (ask.fm)

Aa olin sillee et ||WATAFAK||_{RD}.
 INTERJ be.PST.1SG thus COMP what.a.fuck
 ‘Aa, I was like (lit. **was thus that**) ||WATAFAK||_{RD}.’

(7) Estonian

a. (NMS)

... isa oli körval nii et nagu ||mis sa siis ikka
 father be.PST.3SG nearby so COMP like what 2SG then still
 kihutad||_{RD} ...
 rush.PRS.2SG

‘... father **was** next to me like (lit. **so that like**) ||why are you still rushing||_{RD} ...’

b. (Müürileht)

... ma olin nii, et ||“oh my god, pärisele!”||_{RD}
 1SG be.PST.1SG so COMP INTERJ my god seriously
 ‘... I was like (lit. was so that) ||“oh my god, seriously!”||_{RD}’

The substitution of a reportative verb with a ‘be’-verb frequently leads to the loss of difference between quotations of speech and thought. The equational verb does not specify the type of event behind the RD and only establishes a predicative structure in the QI. I label this process as “event-neutralisation”, i.e. a process which permits different interpretations of the presented quote. In (6a) from Finnish, the type of reported event behind the RD can be deduced from the sequence of two quotes belonging to two different speakers. It is unlikely that the reporter first quotes another speaker’s thoughts and then his/her own thoughts. Instead, (s)he is most likely to represent a dialogue between him-/herself and another speaker. In (7a) from Estonian, the reporter represents the words his father addressed to him, i.e. a quotation of speech. In contrast, (6b) and (7b) illustrate self-quotations that are not explicitly assigned to a concrete addressee. Here the supporting context does not help to distinguish between a quotation of speech or thought.

An even higher degree of event-neutralisation can be observed in Estonian, which allows for the ellipsis of the NP encoding the original speaker, as in (8). Thus, instead of an actual utterance produced by some speaker in a different setting, the reporter presents a quote that (s)he considers emblematic for the described circumstances. The omission of the NP expressing the author of the RD makes this reading the most obvious one.

(8) Estonian (NMS)

Tegelt pärnus on nii et ||vabandust aga ma unustasin
 basically PN.INE be.PRS.3SG so COMP sorry but 1SG forget.PST.1SG
 oma piletit koju||_{RD} ...
 own ticket.PTV home.ILL

‘Basically, in Pärnu it’s (lit. it’s so that) ||sorry, but I forgot my ticket at home||_{RD} ...’

Even though homomorphic constructions do not appear in Finnish, hypothetical quotes can also be introduced by *sillee(n)*. In (9), *sillee(n)* co-occurs with the noun *viesti* ‘message’, encoding the source of the RD. Despite the structural differences between (8) and (9), in both cases the reporters enact fictional discourse through a demonstration for dramatic purposes.

(9) Finnish (Kaksplus)

... laittaa kavereille viestin sillee et ||by the way mulle
put.PRS.3SG friend.PL.ALL message.GEN thus COMP by the way 1SG.ALL
synty viime yönä tyttö / poika mitoilla se ja
be.born.PST.3SG last night.ESS girl boy measure.PL.ADE DEM and
se||_{RD} jne.
DEM etc.

'... sends the message to friends saying (lit. thus that) ||by the way last
night I gave birth to a girl/boy who measures this and that,||_{RD} etc.'

In addition to representations of factual and fictional quotes depicting verbal or mental processes, *sillee(n)* introduces a mimetic expression (10). Besides the representation of enacted human verbal behavior, QIs can also introduce non-linguistic sound imitations, representational gestures or ideophones (Güldemann 2008: 275–295); see §4.3 for a similar instance in Hungarian. The mimetic expression in (10) can be interpreted as a quasi-quotation, where instead of using verbal means, the reporter expresses his/her surprise with the emoticon *o.o* depicting eyes wide open and two question marks.

(10) Finnish (Demi)

Mä olin silleen et ||0.o??||_{RD}
1SG be.PST.1SG thus COMP MIM.MIR
'I was (lit. was thus that) ||0.o??||_{RD}'

Notably, similar functions can be observed online for manner deictics in other languages. See, for example, the German manner deictic *so* in (11), which is used with the same emoticon expressing the reporter's surprise.

(11) German (XHardware)

Ich so ... ||0.o||_{RD} ...
1SG SO MIM.MIR
'I was (lit. I so) ... ||0.o||_{RD} ...'

To sum up, although the Finnish manner deictics can appear in constructions where their Estonian counterpart *nii* is not observed, and vice versa, all function as cataphoric markers. The Finnish manner deictic *niin* is particularly interesting in this regard since it was previously reported to be anaphoric (König 2017: 160). My data shows, however, that *niin* is exclusively attested as a cataphoric quotative marker, and in this it does not differ from the other Finnish manner deictics used in homomorphic constructions.

Although the majority of the examples given here clearly point to a preference for the use of QIs with manner deictics in direct RD, they can also introduce indirect RD. In the non-standard written data considered here, direct RD is usually preferred over indirect. Speakers online often attempt to demonstrate their own or someone else's previously produced or fictional speech and thought rather than to simply describe them. However, as shown for Hungarian in §4.3, the preference for direct or indirect RD can be a crucial factor in the choice of manner deictics in quotative constructions. Table 2 summarises the use of manner deictics in quotative constructions in Finnish and Estonian.

Table 2: Manner deictics in QIs in Finnish and Estonian

Reference	Introduction of MIM expression	Event neutralisation
Finnish		
<i>näin</i>	cataphoric	no + equational verb (marginal)
<i>niin</i>	cataphoric	no
<i>sillee(n)</i>	cataphoric	yes + equational verb
Estonian		
<i>nii</i>	cataphoric	no + equational verb

4.2 Manner deictics in quotative indexes in Permic

Komi and Udmurt have a basic bipartite system of manner deictics: proximal vs. distal. In Udmurt, proximal *tažy* and distal *ožy* do not only indicate a spatial contrast in exophoric function but also have a specific anaphoric function. According to Svetlana Edygarova (p.c.), proximal *tažy* in Udmurt discourse is associated with new information and occupies a pre-focused position. In contrast, distal *ožy* mainly refers to already known or previously mentioned information. In addition, it appears in several idiomatic constructions, e.g. the anaphoric expression *vot ožy* ‘so it is, so it goes’, or as a confirmative particle meaning ‘yes’. Hence, a similar functional division in the quotative domain is expected.

As for Komi, proximal *tadž(i)*¹⁴ and distal *sidž* are formally distinguished and based on the same stems as other demonstratives, e.g. *tajö* ‘this’, *sijö* ‘that, (s)he’. However, they preserve little functional difference in contemporary language, if

¹⁴According to the Komi-Russian dictionary (Beznosikova et al. 2000), the form *tadž* is a shortened variant of *tadži*.

any at all, aside from their distribution in several lexicalised expressions, e.g. *sidž boštny* ‘to take for free (lit. thus)’, *kydž tadž* ‘why did it happen this way (lit. how so)’.¹⁵ Previously, there seemed to be a dialectal difference in the distribution of manner deictics among Komi dialects. In Uotila’s text collections depicting language use from the first half of the 20th century, proximal *tadž(i)* does not appear at all in dialects of the Komi-Permyak literary standard (Uotila 1985); all recorded speakers only use the distal counterpart *sidž*. In dialects of the Komi-Zyrian literary standard (Uotila 1989), one can similarly observe a preference for distal *sidž*. It is used predominantly anaphorically, referring to the manner or events previously described in discourse. Proximal *tadž(i)* is used only once cataphorically; in all other instances, although they are not numerous, it appears in the anaphoric function, similarly to distal *sidž*. Besides these basic manner deictics, Komi also has focused markers that are formed with the prefixed particle *e-* attached to the basic stem of manner deictics: *etadž*, *esidž*. Fedjunëva (2009) argues that focused manner deictics appeared in the language as a relatively recent innovation under the influence of Russian, e.g. *tot* ‘that one’ vs. *è-tot* ‘this one’, *tak* ‘so’ vs. *è-tak* ‘thus (in contrast to so)’. This claim is supported by the lack of corresponding markers in Udmurt and by similarities in the use of demonstratives with the *e*-element in Russian and Komi (Fedjunëva 2009: 95–96).

In quotative constructions of both Udmurt and Komi, manner deictics are used as additional elements in combination with reportative verbs. The focused deictics of Komi are not used in QIs.

In Udmurt quotative constructions, the above-mentioned distinction between the proximal and distal form can also be observed. Most commonly, proximal *tažy* is used cataphorically (12a) in QIs preceding RD, while *ožy* appears as an anaphoric marker following the quote (12b).

(12) Udmurt

- a. (vk.com/udmurt_ept)
Veraške tažy čömyś: “Esli ty ne na Internete ...”||RD
say.PRS.3SG so often {if} {2SG} {NEG} {on} {internet.PREP}
'He often says (lit. says so): "If you are not on the internet ..."||RD'¹⁶
- b. (Blog subcorpus)
“Mon pićiges!”||RD – ačiz šaryś ožy vera.
1SG small.COMPAR self.3SG about thus say.PRS.3SG
'“I am smaller!”||RD – she says (lit. thus she says) about herself.'

¹⁵An identical idiomatic expression is found in Russian, *kak tak*, which is probably the source of the Komi expression.

¹⁶The quote depicts Russian speech produced by a non-native speaker.

However, one can also find instances where the manner deictics are used in the opposite way; consider the anaphoric use of proximal *tažy* in (13).

- (13) Udmurt (*vk.com/udmurt_ept*)
 “Anaj, nu ešče čut’-čut’ ...”||_{RD} – **tažy šuysal** dyr mon ...
 mom {PTCL} {more} {a.bit} so say.COND.1SG maybe 1SG
 ‘“Mom, give me a bit of time [and I will get up, put a kettle on the stove, please]”||_{RD} – that’s what I would probably say (lit. **so I would say** maybe) [if I was at home].’

Example (13) reflects a less systematic use of manner deictics. An investigation of the speech of the blogger on *vk.com/udmurt_ept* also shows that he deviates from the established use of markers outside the quotative domain. In his speech (and that of few other speakers), proximal *tažy* (rather than distal *ožy*) appears in anaphoric reference (14). Furthermore, proximal *tažy* is used instead of distal *ožy* in the anaphoric expression *vot ožy* ‘so it goes’, which is actually a fixed idiomatic expression in contemporary Udmurt.

- (14) Udmurt (*vk.com/udmurt_ept*)
 mon ebašil školae 11 ar i **tažy** mon
 1SG {fucking.go.PST.M} school.ILL 11 year {and} so 1SG
 otmečaju soje bydtemme!???
 {celebrate.PRS.1SG} DEM.ACC end.PTCP.ACC1SG
 ‘[We were sitting and drinking tea with classmates. Walked around a bit. We were watching movies all night. (...)] I was fucking going to the school for 11 years and **so** I celebrate it, my graduation!???’

Similarly, the distal *ožy* is attested as a cataphoric marker pointing at the following quote (15).

- (15) Udmurt (Blog subcorpus)
 a. Vyny tužo maly ke **ožy šuiz:** ||“Oh, mar
 younger.brother.1SG also why INDEF thus say.PST.3SG INTERJ what
 ke so tuž kurdyt ...”||_{RD}
 INDEF DEM very scary
 ‘My younger brother also for some reason said (lit. **said thus**): ||“Wow, she’s somewhat very scary ...”||_{RD}’

- b. Vitalij Agabajev šaryś ozy šusyal: ||so odigez geńialnoj
PN PN about thus say.COND.1SG 3SG one.3SG genius.ADJ
artist val.||_{RD}
artist be.PST.3SG
- 'I could say this (lit. say thus) about Vitaliy Agabaev: ||he was one of
the genius artists.||_{RD}'

Two different explanations can be proposed for the appearance of *ozy* in preposed QIs instead of in the postposed position, which can be considered more common. One scenario suggests that *ozy* is still anaphoric and refers to information previously mentioned in the context and later repeated as a quote. In (15a), this claim may be supported by the presence of the adverb *tužo* 'also'. Thus, one could assume that the reporter quotes her brother, anaphorically referring to an identical utterance that has been previously produced by another speaker in a different context, i.e. 'my brother made an utterance identical/similar to someone else's'; see (16) for the collocation of *ozy* and the focus particle *ik* in a similar context. With respect to (15b), one could assume that the reporter refers to an already familiar opinion about the famous artist. However, in both cases the context does not explicitly support this explanation. Alternatively, (15) may simply present a less systematic use of *ozy*. Thus, similar to proximal *tažy* (13), some idiolects may reflect asymmetry with its more conventionalised use as an anaphoric marker.

Separately, one can also observe the collocation of *ozy* with the particle *ik* in preposed QIs (16). The particle *ik* functions as a focus particle or as a marker signalling the repetition of an element of a situation; cf. the label *marker povtora elementa situacii* in Zubova (2016: 445) and the translation 'the same' in Arkhangelskiy (2014). According to Zubova (2016: 445–446), in the Beserman dialect of Udmurt, the particle *ik* is frequently in collocation with anaphoric elements. In QIs, the collocation of *ozy* and *ik* is used to present RDs already mentioned in a different form in discourse. Thus, in (16) the reporter confirms that somebody was indeed chased away.

- (16) Udmurt (vk.com/knyazpozdey)
Ozy ik šuizy: ||"Myn tatyś!"||_{RD}
thus PTCL say.PST.3PL go.IMP.2SG here.ELA
'They also said (lit. said thus): ||"Go away from here!"||_{RD}'

In sum, in contemporary Udmurt, the manner deictics are used in quotative constructions as follows: proximal *tažy* is primarily a cataphoric marker, refer-

ring to following quotes; distal *ožy* is mainly used anaphorically following quotes. The appearance of the distal manner deictic *ožy* as a part of a preposed QI is quite common when *ožy* collocates with the additive particle *ik*. Thus, *ožy ik* serves as a reference to a previously described event repeated as a quote.

Some less frequent uses of the manner deictics that do not fall under the above rules may reflect the decreasing linguistic intuition of Udmurt speakers under the influence of Russian, which uses the proximal manner deictic *tak* both for anaphoric and cataphoric reference (17). The distal (*sjak*) and contrastive manner deictic (*ètak*) do not appear systematically in Russian QIs.

(17) Russian

a. (Woman.ru)

Êto on tak skazal, čto ||žit' v meste budete||_{RD} ...
DEM he so say.PST.M COMP live.INF together be.FUT.2PL

'This is him who said (lit. said so that) ||you will live together||_{RD} ...'

b. (Vremja Novostej Online)

||"Neobxodimo doždat'sja kopii protokola ..."||_{RD} tak
necessary wait.INF copy.GEN record.GEN so
otvetil na vopros ... Geral'd [sic!] Jaroš.
answer.PST.M on question PN PN

||"It's necessary to wait for the copy of the record [and then conduct corresponding steps]"||_{RD} – so answered Gerald [sic!] Jarosch to a question ...'

In Komi quotative constructions with manner deictics, it is observed that only the proximal marker *tadž(i)* 'so' appears with reportative verbs, (18a) and (18c), and inchoative verbs (18b) in forming QIs. In the data, such combinations are not frequent and account for only three instances. However, this figure should not be taken as definitive and may only show less frequent quotative use of manner deictics in Komi new media texts. It is noteworthy that the proximal marker only appears in Komi QIs in internet communications, whereas distal *sidž* 'thus' can be used in quotative constructions outside the new media genre (19).

(18) Komi

a. (Alëna Tuvsovja)

... mövpyšti tadž: ||"Metög na udžalyšjas šuröny ..."||_{RD}
think.PST.1SG so 1SG.ABE PTCL worker.PL get.found.PRS.3PL

'[Like most of the people,] I thought (lit. thought so): ||"The workers will be found without me ..."||_{RD}'

b. (vk.com/biarmian)

A zavoditčö tadži: ||“Sovetskøj obščestvennost’ ...”||_{RD}

and begin.PRS.3SG so Soviet {community}

‘And it begins (lit. **begins so**): ||“The Soviet community ...”||_{RD},

c. (Tuš)

||“Nevažno v kakom krutom vuze
unimportant.ADV} {in} {which.PREP} {cool.PREP} {university.PREP}

ty učiš’sja, važno sumet’ sebja

{2SG} {study.PRS.2SG} {important.ADV} {manage.INF} {self.ACC}

realizovat”||_{RD} 11-öd klassyn velödčigön tadži menym
{fulfill.INF} 11-ORD class.INE study.CV.INSTR SO 1SG.DAT

vištalis gimnazijasa radejtana velödyşjasyś öfi Alla
tell.PST.3SG gymnasium.ADJ favorite teacher.PL.ELA one PN

Aleksandrovna Taskajeva.

PN PN

‘It isn’t important what kind of cool university you are studying at,
it is important to manage to fulfil yourself;’||_{RD} so I was told (lit. **so
told me**) when I studied in the 11th class of one of the gymnasium’s
favourite teachers, [the one (called)] Alla Aleksandrovna Taskayeva.’

Since *tadž(i)* is only rarely employed in my data, I turned to the available text collections (Uotila 1985; 1989) and the digital corpus that assembles texts not belonging to the new media genre (see §3) to see whether they show different results. As mentioned at the beginning of this subsection, in Komi-Permyak dialects (Uotila 1985), only distal *sidž* is used. It also appears in quotative constructions referring cataphorically to the following quote (19). In Komi-Zyrian texts (Uotila 1989), both markers are used but neither appears in quotative constructions.

(19) Komi-Permyak (Uotila 1985: 40; glossing and translation mine)

a sar’ vištalis sidž: ||“on-kö addžy čuńkyčlö, me
and tzar tell.PST.3SG thus NEG.2SG-PTCL.COND see.CN ring.DAT 1SG
tenö vija.”||_{RD}
2SG.ACC kill.PRS.1SG

‘And the tzar said (lit. **said thus**): ||“If you don’t find the ring, I will kill
you.”||_{RD}’

Available texts from the digital corpus show that both manner deictics are used in QIs. However, distal *sidž* is more frequent than proximal *tadž(i)*. As for the

referential function, both markers can be used either anaphorically or cataphorically, as in (20). Distal *sidž* is more common in cataphoric functions (20a); it is only attested twice in anaphoric functions (20b). In contrast, proximal *tadž(i)* is predominantly used anaphorically (20c); only some instances reflect a cataphoric use (20d).

(20) Komi (Komi corpus)

- a. ... i **sidž šuöny:** ||byťfökö seni sijö götyr pydđi olö.||_{RD}
and thus say.PRS.3PL as.if there 3SG wife instead live.PRS.3SG
'... they say (lit. say thus): ||as if she lives there instead of the wife.||_{RD}'
- b. ||A menym kolö koršny arlyda ñin
but 1SG.DAT must.PRS.3SG find.INF in.years.ADJ already
mortös||_{RD} – **sidž dumajtis** Nikolaj ...
person.ACC1SG thus think.PST.3SG PN
'But I have to find a person who is already elderly,||_{RD} – thought (lit.
thus thought) Nikolay ...'
- c. ||“Bur olömsö oškyšta, (...)”||_{RD} – **tadži šuö** poet.
good life.ACC.3SG praise.PRS.1SG so say.PRS.3SG poet
'“Good life I praise, [bad life I wipe away],”||_{RD} – **so said** the poet.'
- d. Tajö sijö **tadži šuö:** ||“Nyvka, a komandovajtö!”||_{RD}
DEM.PROX 3SG so say.PRS.3SG girl but give.orders.PRS.3SG
'He said (lit. said it so): ||“A girl, but she gives orders!”||_{RD}'

Despite the lack of a representative number of examples in my corpus, other materials show that manner deictics can refer indiscriminately to preceding and following RDs in contemporary Komi QIs. This use of manner deictics resembles the use of *tak* in Russian and could be motivated by Russian influence. Different Komi speakers may associate different autochthonous markers with Russian *tak* (17) and use them according to the Russian model. Such pattern replications are frequently observed in contemporary Komi as a way to preserve Komi language use through the choice of autochthonous markers in constructions modelled on Russian, or of language features that could be common to both languages (see Leinonen 2006; 2009). Thus, Komi-Permyak speakers may turn to the distal marker as the closest equivalent to Russian *tak*, while in Komi-Zyrian one can observe the use of distal *sidž* more frequently; note, however, that proximal *tadž(i)* also appears as the only option in several idiolects, as reflected in my data (18).

The use of Udmurt and Komi manner deictics is summarised in Table 3.

Table 3: Manner deictics in QIs in Udmurt and Komi

Primary endophoric meaning		Reference	Position of QIs
Udmurt			
<i>tažy</i>	new information	primarily cataphoric	pre- and post-posed
<i>ožy</i>	known information	primarily anaphoric	pre- and post-posed
Komi			
<i>tadž(i)</i>	no specialisation	cata- and anaphoric	pre- and post-posed
<i>sidž</i>	no specialisation	cata- and anaphoric	pre- and post-posed

4.3 Manner deictics in quotative indexes in Hungarian

Hungarian has a bipartite system of manner deictics, contrasting proximal *így* (1) and distal *úgy*. These markers have adopted distinct functions in the quotative domain. Proximal *így* is used as a general mimetic marker introducing a demonstration into discourse (Beáta Gyuris, p.c.) (21a). In turn, distal *úgy* functions as a marker pointing at the manner of action. Hence, instead of demonstrations, the marker is usually followed by a comparison with another action (Beáta Gyuris, p.c.), as in (21b).¹⁷

(21) Hungarian

- a. (Beáta Gyuris, p.c.)
Így csinálom: {demonstration}.
 so do.PRS.1SG.DEF
 'I do it like this: {demonstration}'.
- b. (gyakorikeredesek.hu)
Én is úgy csinálom mint te.
 1SG also thus do.PRS.1SG.DEF like 2SG
 'I also do it like you.'

Similarly to Udmurt, the functions of the manner deictics outside the quotative domain affect their use in quotative constructions. This effect is reflected in

¹⁷I am not aware of studies addressing this functional division of manner deictics in contemporary Hungarian.

two aspects: (i) the types of RD, and (ii) the position of the QI within the RD-construction.

As a mimetic marker, *így* is used only with direct RD that “reports and demonstrates what has been uttered” (Dömötör 2001: 338). In mimetic expressions, demonstrations are carried out by movements and gestures. In quotative constructions, RD is a demonstration of somebody’s words (22a) or mental activities (22b) (see Clark & Gerrig 1990 on quotations as demonstrations). Note that, in Hungarian, even direct RD can be preceded by the complementiser *hogy* (22c). Thus, unlike in many SAE languages, the presence of the complementiser (22b) is not a sufficient criterion for distinguishing between indirect and direct RD.

(22) Hungarian

- a. (MNSz)

a rapper rövid beszédében **így mutatta** **be magát:**
 DEF rapper short talk.3SG.INE SO show.PRS.3SG.DEF into self.ACC
 ||“cigány vagyok, zsidó vagyok ...”||_{RD}
 Gypsy be.PRS.1SG Jew be.PRS.1SG

‘The rapper **introduced** himself in his short speech by saying (lit. **so**):
 ||“I am Gypsy, I am Jew ...”||_{RD}’
- b. (MNSz)

Én is **így gondoltam** **hogy** ||nem túl jó választás nekem
 1SG also so think.PST.1SG COMP NEG very good choice DAT.1SG
 a 29er ...||_{RD}
 DEF 29er

‘I also thought (lit. **thought it so that**) ||the 29er is not a very good choice for me ...||_{RD}’
- c. (Google books)

... azt **mondata**, **hogy** ||szeretlek, drága
 DEM.ACC say.PST.3SG.DEF COMP love.PRS.1SG.DEF.2SG expensive
 csillagom.||_{RD}
 star.1SG

‘... (he) said (lit. **said that**) ||I love you, my dear star.||_{RD}’

As in Finnish (10), some of the Hungarian mimetic expressions introduced by a QI with *így* can be interpreted as quasi-quotations. In (23), the reporter puts a person’s moves and gestures into words.

(23) Hungarian (MNSz)

- ... a kezével így mutatta, hogy ||ha nem, akkor ez a
DEF hand.3sg.COM so show.PST.3SG.DEF COMP if NEG then DEM DEF
táská bumm.||_{RD}
bag IDEO
'... with his hand he showed (lit. showed so) that ||if not, then this bag will
go boom.'||_{RD}

Functionally, distal *úgy* introduces the content of a proposition expressed as a quote, rather than a mimetic quote. Thus, QIs with *úgy* do not show any restriction to one type of RD and can be used with either direct (24a) or indirect (24b) quotations.

(24) Hungarian

a. (MNSz)

- ... úgy mondta: ||látom nagyon istenfélők
thus say.PST.3SG.DEF see.PRS.1SG.DEF very God.fearing.PL
vagytok||_{RD} ...
be.PRS.2PL
'... said (lit. said it thus): ||I see you are very God-fearing'||_{RD} ...'

b. (nepmese.hu)

- Úgy mondta, hogy ||kell neki a száz
thus say.PRS.3SG.DEF COMP need.PRS.3SG DAT.3SG DEF 100
forint ...||_{RD}
forint
'[He] thus said that ||he needs the one hundred forints'||_{RD} ...'

In principle, indirect quotations as in (24b) do not contradict the statement that *úgy* introduces only the content of a quote, differing from the original utterance at least in deictic orientation, i.e. 'he needs the one hundred forints' vs. 'I need the one hundred forints'. In addition, one can expect that the choice of original words might have differed from those reported, e.g. 'Give me, please, one hundred forints' or 'Could you lend me the one hundred forints'. As Dömötör (2001: 338) points out, "[w]hile direct speech [equivalent to direct RD] reports and demonstrates what has been uttered, indirect speech [equivalent to indirect RD] renounces this demonstration (...) [indicating] that the author has understood the utterance and based on this, he reformulates its content". Yet, this

statement is problematic with direct quotations such as (24a), which resemble a demonstration of a person's utterance rather than a mere depiction of its content. A closer look at the syntactic distribution of the manner deictics reveals that the proximal *így* is not restricted to one position within the RD-construction: it can appear preceding (22)–(23) or following (25a) the RD, or it can split the RD into parts (25b). In contrast, QIs with the distal *úgy* are restricted to the pre-RD position.

(25) Hungarian (MNSz)

- a. ||De egy tehénnel tette||_{RD} – így a riporter.
but INDEF cow.COM do.PRS.3SG.DEF so DEF reporter
‘But he did it with a cow||_{RD} – so [said] the reporter.’
- b. ||Á nem,||_{RD} (így ő), ||hisz mi immár kétezer
INTERJ NEG so 3SG believe.PRS.3SG 1PL now two.thousand
éve imádkozunk Jeruzsálemért.||_{RD}
year.3SG pray.PRS.1PL Jerusalem.CAUS
‘Ah no,||_{RD} (so he [said]), ||we have already prayed for Jerusalem for
two thousand years.||_{RD}’

As shown in (25), QIs with the proximal *így* do not always contain a speech or epistemic verb. Event-neutralised QIs (§4.1) consist merely of an NP referring to the original speaker and the manner deictic, as in (25a) and (25b). In contrast, QIs with the distal *úgy* must contain reportative verbs, otherwise the whole construction is considered ungrammatical. Hence, in the case of proximal *így*, the RD-construction can be considered a subtype of mimetic construction, i.e. ‘X is/was like this: {demonstration}’, while in the case of distal *úgy*, the RD-construction is a reproduction of somebody’s words or thoughts: ‘X says/thinks something like that: {quote}’.

Previous studies have already pointed out the additional meanings expressed by the manner deictics in Hungarian quotative constructions. Kiefer (2016: 83) indicates that “if the exact wording of the reported utterance is at stake, the reported utterance is repeated but the reporting clause contains the [proximal] adverbial particle *így* ‘so, thus’”. Körtvély (2016: 607), in turn, mentions that “[n]ative speakers of Hungarian estimate both versions [of constructions with and without the distal *úgy*] as quasi equivalent in their meaning; however, some of them consider the contents of the *úgy*-type as less certain”.

To complement Kiefer’s statement, my investigation shows that QIs with *így* can also introduce hypothetical quotes, as in (26), where the reporter presents a

quote purportedly said by the world-famous Hungarian football player, Ferenc Puskás.

(26) Hungarian (MNSz)

“Nagy pénz, kis foci.”_{RD} – vagy mégsem így mondta
big money small football or still.also.NEG SO say.PST.3SG.DEF
volna Puskás?
be.COND.3SG PN

‘“Big money, small football.”_{RD} – as Puskás would have said (lit. or
wouldn’t have Puskás said so)?’

Note that Puskás’s actual utterance *Kis pénz – kis foci, nagy pénz – nagy foci* ‘Small money – small football, big money – big football’,¹⁸ differs from the one presented in (26). Thus, instead of presenting the original utterance, the reporter modifies it according to his/her aims. Since *így* can introduce hypothetical quotes, I suggest revising Kiefer’s claim: instead of associating the use of the proximal manner deictic *így* with the exactness of the presented quote, I propose associating it with the directness of the quote. Of course, direct quotes introduced by the proximal manner deictic may be represented by verbatim quotations. However, this condition is not necessarily true for the whole category of direct RD.

As for Körtvély’s claim, it is to be expected that indirect quotes introduced by QIs with *úgy* can be perceived as approximately reproduced and signal the reporter’s lack of commitment to the content of quote. First, *úgy* introduces the content of quote only, which leads to the difference between the original utterance and its reproduction. Second, several features of RD might be blurred due to syntactic adjustment, while presenting indirect RD. Although both conditions are found in the use of *úgy*, the reporter’s lack of commitment is hardly foregrounded in all RD-constructions where the distal manner deictic is used. As a result, I step back from the analysis proposed above and propose considering both meanings assigned for the manner deictics secondary, rather than universally applicable to their use in the quotative domain.

The main features drawing differences between proximal *így* and distal *úgy* are summarised in Table 4.

¹⁸This utterance is frequently attributed to Puskás, although there is no actual proof that he has ever said it.

Table 4: Features of manner deictics in Hungarian QIs

	Proximal <i>így</i>	Distal <i>úgy</i>
Primary meaning	demonstration of RD	representation of content of RD
RD type	direct	direct, indirect
Reference	cata- and anaphoric	cataphoric
Position of QIs	pre-, intra- and postposed	preposed
Event-neutralisation	ellipsis of the verb	does not occur

5 Summary and discussion

This chapter has shown that manner deictics follow different patterns of distribution in QIs in related languages. According to their referential function, three main types of manner deictics are distinguished: (i) cataphoric, (ii) anaphoric, and (iii) both cataphoric and anaphoric deictics; cf. the summary in Table 5. In addition to their referential function, manner deictics show preferences for the direct or indirect types of RD and for the pre-, intra- or postposed position in the RD-construction.

Table 5: Referential functions of manner deictics in quotative constructions of selected Finno-Ugric languages

Referential functions	Manner deictics
Cataphoric	Finnish <i>näin</i> , <i>niin</i> , <i>sillee(n)</i> Estonian <i>nii</i> Hungarian <i>úgy</i> Udmurt <i>tažy</i>
Anaphoric	Udmurt <i>ožy</i>
Cata- and anaphoric	Komi <i>tadž(i)</i> , <i>sidž</i> Hungarian <i>így</i>

Among the five Finno-Ugric languages, several patterns of distribution can be identified based on the above criteria. In languages with a proximal/distal pair of manner deictics, their meanings and functions outside the quotative domain

predetermine their use in QIs. Thus, the Hungarian proximal manner deictic *íggy* as a general mimetic marker introduces direct RD only. Distal *úgy*, which outside the quotative domain can only refer to the manner of action but not demonstrate it, is used as a marker pointing at the representation of the content of a person's words or thoughts. Therefore, it can introduce both direct and indirect RD.

In Udmurt, the distribution follows from the meanings of the manner deictics. Proximal *tažy* introduces new information and, in RD-constructions, typically appears in preposed QIs. Distal *ožy*, in turn, refers to already-known information and is used in postposed QIs.

In contrast, in Komi, the proximal and distal manner deictics are interchangeable outside the quotative domain. In the quotative domain, their use seems influenced by the genre in which they are used: distal *sidž* is used more frequently in conventional written texts, while in my new media material only proximal *tadž(i)* is attested in quotative constructions.

In Finnish, proximal and distal manner deictics are exclusively used with cataphoric reference in quotative constructions. This is all the more interesting as the distal manner deictic *niin* is also used as an anaphoric marker outside the quotative domain (König 2017: 160). The Estonian cognate *nii* has been shown to be used only as a cataphoric marker, which confirms Keevallik's (2005: 116–117) earlier findings. In colloquial speech, Finnish *sillee(n)* and Estonian *nii* co-occur with reportative verbs or with equational verbs in QIs. QIs with equational verbs can equally introduce quotations of speech and thought, while speech or epistemic verbs restrict the interpretation to one type of RD. Manner deictics can appear in contexts where the RD is not attributed to a concrete speaker and is therefore hypothetical. Usually, the QI introducing hypothetical quotes is structurally less complex.

In addition, two manner deictics, Finnish *sillee(n)* and Hungarian *íggy*, are also observed introducing mimetic expressions that can be interpreted as quasi-quotations. It is not surprising that the demonstratives are employed as quotatives with mimetic expressions. Güldemann (2008: 521) mentions that many quotative markers have been initially used in indexing mimesis and later on were grammaticalised into exclusively quotative markers or are still employed in their initial function parallel to their relatively new quotative use.

The observations made for Finno-Ugric languages have parallels in other languages of the world. For example, an exclusively cataphoric manner deictic is found in Usan (Papuan), where “the quote introduction has the cataphoric adverbial *ete* [*e-t-e* ‘this/her’ + postposition *-t* ‘for/at/on/etc.’ + ‘this/her’]: *wo ete qamar* ‘he said thus’” (Reesink 1993: 218). Among the markers that can be used

both cataphorically and anaphorically in QIs, I have already mentioned the Russian manner deictic *tak*. Similarities can be pointed out in the use of German *so* and French *ainsi*, both meaning ‘so’, that appear preceding and following the RD (König 2017: 160; Karszenberg & Lahousse 2018; also see König this volume and Diessel & Breunesse this volume). Similarly, in Usan QIs, the manner deictic *ende* ‘this/here’ + ‘given/a particular one’ + postposition *-t* ‘for/at/on/etc.’ + ‘this/here’ can be used both cata- and anaphorically (Reesink 1993: 218).

I am not aware of other languages in which the proximal form of a proximal/distal manner deictics pair is used for direct RD/new information, while the distal form is used for indirect RD/old information, as has been shown for Hungarian and Udmurt, respectively. However, one could expect to find similar patterns in quotative systems of languages not considered here, which is a direction for future research.

Acknowledgments

I am grateful to Gerson Klumpp for his insightful comments on an earlier version of this paper. I would also like to thank the two anonymous reviewers and Yvonne Treis for their comments and suggestions, which have greatly improved this chapter. The responsibility for all remaining shortcomings is entirely mine.

Abbreviations

ABE	abessive case	DAT	dative case
ACC	accusative case	DEF	definite
ADE	adessive case	DEM	demonstrative
ADJ	adjective	ELA	elative case
ADV	adverb	ESS	essive case
ALL	allative case	FUT	future tense
CAUS	causative case	GEN	genitive case
CMC	computer-mediated communication	IDEO	ideophone
CN	connegative	ILL	illative case
COM	comitative case	IMP	imperative mood
COMP	complementiser	INDEF	indefinite
COMPAR	comparative	INE	inessive case
COND	conditional mood	INF	infinitive
CV	converb	INSTR	instrumental case
		INTERJ	interjection

M	masculine gender	PRS	present tense
MIM	mimetic	PST	past tense
MIR	mirative	PTCL	particle
NEG	negative	PTCP	participle
ORD	ordinal	PTV	partitive case
PROX	proximal	QI	quotative index
PL	plural	QP	question particle
PN	proper noun	RD	reported discourse
PP	past participle	SG	singular
PREP	prepositional case	SNS	social network sites

Data sources

The data sources were last accessed in the period from 1 February to 1 August 2019.

Electronic corpora

Blog subcorpus = Blog subcorpus of Udmurt corpus:

<http://web-corpora.net/UdmurtCorpus/search/>

Internet-keskusteluainestojä [The databases of internet communications]:

<https://korp.csc.fi>

Komi corpus = Korpus komi jazyka [The corpus of the Komi language]:

<http://komicorpora.ru>

MNSz = Magyar Nemzeti Szövegtár. Személyes alkorporusz [The Hungarian National Corpus. Personal Subcorpus]:

<http://corpus.nytud.hu/mnsz/>

NMS = Eesti keele koondkorpus. Uus media [Estonian Reference Corpus. New media subcorpus]:

<https://www.cl.ut.ee/korpused/segakorpus/uusmeedia/>

Internet sources

Alëna Tuvsovja:

http://tuvsovja.blogspot.com/2014/09/blog-post_6.html

ask.fm

https://ask.fm/Joonas_Tuloneen/answers/120680473973

Daily Mail

<https://www.dailymail.co.uk/news/article-3030136/Abbott-deserved-credit-McGregor.html>

Demi

<https://www.demi.fi/keskustelu/suhteet-aaah-mita-ma-teen>;
<https://www.demi.fi/keskustelu/suhteet-miten-unohan-jatkan>

Eestimaa Loomakaitse Liit

<http://loomakaitse.eu/lapsed-kui-te-seda-kassipoega-ei-vota-siis-ma-tapan-tara/>

Google books

<https://books.google.ee/books?id=xQbWDwAAQBAJ&lpg=PT415&dq=%22azt%20mondta%20hogy%20szeretlek%20dr%C3%A1ga%20csillagom%22&pg=PP1#v=onepage&q&f=false>

gyakorikerdesek.hu

https://www.gyakorikerdesek.hu/gyerekvallalas-neveles_babak_1253137-halozsak-

Kaksplus

<https://kaksplus.fi/threads/alkionsiirroista-plussanneet.2458857/page-11>

lansiuusimaa.fi

<https://www.tallinna24.ee/blogit/satunnainen-herrasmies>

Müürileht

<https://www.muurileht.ee/skoda-mazda-ja-ritmos-liquidos/>

nepmese.hu

<https://www.nepmese.hu/index.php/mesetar/mesek/adj-uram-isten-szaz-forintot>

Tuš

<http://tusjuk.blogspot.com/2015/>

Twitter

<https://twitter.com/allthingsmollie/status/690042054936772608>

vk.com/biarmian

<https://vk.com/biarmian>

vk.com/knyazpozdey

<https://vk.com/knyazpozdey>

vk.com/udmurt_ept

https://vk.com/udmurt_ept

Vremja Novostej Online [Time of News Online]

<http://www.vremya.ru/2009/74/5/228155.html>

Woman.ru

<http://www.woman.ru/relations/men/thread/4996059/>

XHardware

<http://extreme.pcgameshardware.de/mainboards-und-arbeitsspeicher/257287-ram-speed.html>

References

- Arkhangelskiy, Timofey. 2014. *Clitics in the Beserman dialect of Udmurt* (Basic Research Program Working Papers, WP BRP 10/LNG). Moscow: National Research University Higher School of Economics.
- Beznosikova, Ljutsija M., Evgenija A. Ajbabina & Rajsa I. Kosnyreva. 2000. *Komi-roč kyvčukör* [Komi-Russion dictionary]. Syktyvkar: Komi knižnoe izdatel'stvo.
- Buchstaller, Isabelle & Ingrid van Alphen. 2012. Preface: Introductory remarks on new and old quotatives. In Isabelle Buchstaller & Ingrid van Alphen (eds.), *Quotatives: Cross-linguistic and cross-disciplinary perspectives*, xii–xxx. Amsterdam: Benjamins.
- Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123(3). 324–347.
- Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805.
- Crystal, David. 2001. *Language and the internet*. Cambridge: Cambridge University Press.
- Dömötör, Adrienne. 2001. Tendencies in the development of Late Old Hungarian and Early Middle Hungarian main clauses of reported speech. *Acta Linguistica Hungarica* 48(4). 337–369.
- Edygarova, Svetlana. 2013. Ob osnovnyx raznovidnostjax sovremennoj udmurtskogo jazyka [On fundamental varieties of the contemporary Udmurt language]. *Ežegodnik Finno-Ugorskix Issledovanij* 3. 7–18.
- Edygarova, Svetlana. 2014. The varieties of the modern Udmurt language. *Finnisch-Ugrische Forschungen* 62. 376–398.
- Erelt, Mati & Helle Metslang (eds.). 2017. *Eesti keele sūntaks* [The syntax of Estonian]. Tartu: Tartu Ülikooli Kirjastus.

- Fedjunëva, Galina V. 2009. O refleksax praural'skix dejkičeskix častic *e 'etot, tot' i *o 'tot' v permskix jazykax [On the reflexes of the Proto-Uralic deictic particles *e 'this, that' and *o 'that' in Permic languages]. *Voprosy Jazykoznanija* 2009(1). 91–97.
- Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.
- Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen & Irja Alho. 2004. *Iso suomen kielioppi* [The large grammar of Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Helasvuo, Marja-Liisa, Marjut Johansson & Sanna-Kaisa Tanskanen. 2014. Johdatus digitaalisen vuorovaikutukseen [Introduction to digital interaction]. In *Kieli verkossa: Näkökulmia digitaaliseen vuorovaikutukseen* [Language online: Perspectives on digital interaction], 9–20. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Karssenberg, Lena & Karen Lahousse. 2018. On the different interpretations of sentence-initial *ainsi* 'so' and the competition between three types of verb-subject order. *Folia Linguistica* 52(1). 1–38.
- Keevallik, Leelo. 2005. The deictic *nii* 'so, in this way' in interaction. In Daniele Monticelli, Renate Pajusalu & Anu Treikelder (eds.), *From utterance to uttering and vice versa: Multidisciplinary views on deixis*, 109–126. Tartu: Tartu Ülikooli Kirjastus.
- Kel'makov, Valentin K. 1981. *Obrazcy udmurtskoj reči: Severnoe narečie i sredinnye govory* [Examples of Udmurt speech. Northern dialect and central varieties]. Iževsk: Naučno-issledovatel'skij institut pri sovete ministrov Udmurtskoj ASSR.
- Kel'makov, Valentin K. 1990. *Obrazcy udmurtskoj reči II: Sredinnye govory* [Examples of Udmurt speech II: Central varieties]. Iževsk: Udmurtija.
- Kiefer, Ferenc. 2016. Indirect and direct reports in Hungarian. In Alessandro Capone, Ferenc Kiefer & Franco Lo Piparo (eds.), *Indirect reports and pragmatics: Interdisciplinary studies*, 77–93. Cham: Springer.
- König, Ekkehard. 2017. The deictic identification of similarity. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 143–164. Amsterdam: Benjamins.
- König, Ekkehard & Carla Umbach. 2018. Demonstratives of manner, of quality and of degree: A neglected subclass. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives: Syntax, semantics and pragmatics*, 285–328. Berlin: De Gruyter Mouton.

- Körtvély, Erika. 2016. Modal functions of the complementizer *hogy* ‘that’ in Hungarian. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages*, 587–618. Berlin: Mouton de Gruyter.
- Leinonen, Marja. 2006. The Russification of Komi. *Slavica Helsingiensia* 27. 234–245.
- Leinonen, Marja. 2009. Russian influence on the Ižma Komi dialect. *International Journal of Bilingualism* 13(3). 309–329.
- Pischlöger, Christian. 2014. Udmurtness in Web 2.0: Urban Udmurts resisting language shift. *Finnisch-Ugrische Mitteilungen* 38. 143–162.
- Pischlöger, Christian. 2016. Udmurt on social network sites: A comparison with the Welsh case. In Reetta Toivanen & Janne Saarikivi (eds.), *Linguistic genocide or superdiversity? New and old language diversities*, 108–132. Bristol: Multilingual Matters.
- Reesink, Gerard Pieter. 1993. Inner speech in Papuan languages. *Language and Linguistics in Melanesia* 24. 217–225.
- Teptiuk, Denys. 2019. New quotatives in Finnish and Estonian. *Finnisch-Ugrische Mitteilungen* 42. 207–242.
- Uotila, T. E. 1985. *Syrjänische Texte I: Komi-Permjakisch*. Helsinki: Suomalais-Ugrilainen Seura.
- Uotila, T. E. 1989. *Syrjänische Texte III: Komi-Syrjänisch: Luza-Letka-, Ober-Sysola-, Mittel-Sysola, Prisyktyvkar-, Unter-Vyčegda- und Udora-Dialekte*. Helsinki: Suomalais-Ugrilainen Seura.
- Zubova, Yulia N. 2016. Sintaksis i semantika fokusnyx častic v bessermjanskom dialekte udmurtskogo yazyka (na primere častic *ik* i *uk*) [Syntax and semantics of focus particles in Besermyan Udmurt (case study of particles *ik* and *uk*)]. *Acta Linguistica Petropolitana* 12(1). 442–451.

Chapter 12

A typology of demonstrative clause linkers

Holger Diessel

Friedrich Schiller University Jena

Merlijn Breunesse

University of Amsterdam

Across languages, demonstratives provide a frequent diachronic source for a wide range of grammatical markers, including certain types of clause linkers such as English *so*, *that*, *thus* and *therefore*. Drawing on data from a sample of 100 languages, this chapter presents a cross-linguistic survey of (grammaticalised) demonstratives that are routinely used to combine clauses or propositions. The study shows that demonstrative clause linkers occur in a large variety of constructions including all major types of subordinate clauses and paratactic sentences. Concentrating on the most frequent types, the chapter considers (grammaticalised) demonstratives functioning as (i) relative pronouns, (ii) linking and nominalising articles, (iii) quotative markers, (iv) complementisers, (v) conjunctive adverbs, (vi) adverbial subordinate conjunctions, (vii) correlatives and (viii) topic markers. It is the purpose of the chapter to provide a comprehensive overview of demonstrative clause linkers from a cross-linguistic perspective and to consider the mechanisms of change that are involved in the grammaticalisation of demonstratives in clause linkage constructions.

1 Introduction

Demonstratives are a unique class of expressions that are foundational to social interaction, discourse processing and grammar evolution (Diessel 2006; 2013; 2014). In face-to-face conversation, demonstratives are commonly used with reference to entities in the surrounding speech situation in order to coordinate the



interlocutors' joint focus of attention. In this use, they are often accompanied by pointing gestures and other non-verbal means of deictic communication (Bühler 1934; see also Coventry et al. 2008).

All languages use demonstratives for spatial reference, but demonstratives are also frequently used with reference to linguistic elements in discourse (Halliday & Hasan 1976). Two basic discourse uses are commonly distinguished: the tracking use, in which demonstratives refer to discourse participants, and the discourse-deictic use, in which demonstratives refer to an adjacent clause or proposition (Webber 1991).

In addition to these uses, many languages have grammatical function morphemes that are historically derived from demonstratives. In the grammaticalisation literature, it is often assumed that all grammatical function morphemes are ultimately based on content words (Heine & Kuteva 2007: 111), but, as Brugmann (1904) and Bühler (1934) noted, demonstratives also provide a frequent source for the development of grammatical markers. There is a wide range of grammatical function words that are frequently derived from tracking and discourse deictic demonstratives, including definite articles, third person pronouns, relative pronouns, copulas and subordinate conjunctions (Himmelmann 1997; Diessel 1999a; 1999b). Some of these markers have been studied intensively from both diachronic and cross-linguistic perspectives. There is, for instance, a great deal of research on the development of definite articles from tracking or anaphoric demonstratives in a large number of languages (e.g. Harris 1978; Cyr 1993; Laury 1997). However, other types of development have not been systematically investigated from a cross-linguistic perspective. Conjunctive adverbs, for instance, are frequently based on discourse-deictic demonstratives, but there is almost no research on this topic (Diessel 1999a: 125–127).

In this chapter, we will be concerned with (grammaticalised) demonstratives that are routinely used for clause linkage. In English, for example, the expressions *so*, *that*, *so that*, *thus* and *therefore* are based on demonstratives and serve to combine clauses or propositions. Similar types of grammaticalised demonstratives occur in many other languages (Himmelmann 1997; Heine & Kuteva 2007). It is the purpose of this chapter to show that demonstratives are of central significance to the development of grammatical markers in the domain of clause linkage. More specifically, the chapter provides a typology of “demonstrative clause linkers” and analyses the mechanisms of change behind their development.

Since demonstratives are commonly used with reference to linguistic elements in the unfolding discourse, they provide a natural starting point for the grammaticalisation of clause linkers (Bühler 1934; Diessel 2012). Yet, while the frequent

development of demonstratives into clause-linking morphemes is motivated by discourse-pragmatic factors, it is important to recognise that the grammaticalisation of clause linkers is also influenced by the syntactic properties of demonstratives in particular constructions (Himmelmann 1997; Diessel 1999a; 1999b).

Early research on grammaticalisation has focused on semantic and pragmatic changes of lexical expressions, but more recent research has shown that grammaticalisation processes involve constructions (Traugott 2003), or entire networks of constructions, rather than just isolated items (Traugott & Trousdale 2013; see also Diessel 2019a). Thus, in order to understand how demonstratives grammaticalise into clause linkers, one must not only consider their discourse functions but also their occurrence in particular constructions.

This chapter builds on previous research on the grammaticalisation of demonstratives (e.g. Brugmann 1904; Bühler 1934; Himmelmann 1997; Diessel 1999a; 1999b; 2006; 2014) but is more detailed and comprehensive than all former accounts. In fact, to the best of our knowledge, this is the first large-scale typological study that systematically investigates the role of (grammaticalised) demonstratives in the domain of clause linkage. The analysis is based on a typological database including information from a genetically and geographically dispersed sample of 100 languages. The languages come from 80 genera, with maximally two languages from each genus, and are roughly equally distributed across the six major geographical areas that are commonly distinguished in typology, i.e. Eurasia, Africa, Southeast Asia and Oceania, Australia and New Guinea, South America and North America (Dryer 1992). The bulk of the data have been gleaned from reference grammars and other published sources, but for some languages we also consulted native speakers and language experts. A complete list of languages included in our sample is given in the appendix.

Most of the variables in our database concern parameters of synchronic variation, but we have also gathered information on the diachronic developments of complex sentences and the various types of clause linkers. Since many clause linkers are only weakly grammaticalised, they are (often) etymologically transparent. There is plenty of evidence in our database that relative markers, complementisers, conjunctive adverbs and many other types of clause linkers are etymologically related to demonstratives. However, the mechanisms of change that are involved in the diachronic development of demonstrative clause linkers are often difficult to analyse. As we will see, in many cases we know that a particular clause linker has a deictic origin, but since there are no diachronic corpora to study the constructional changes that are involved in the grammaticalisation of demonstratives into clause linkers, we do not always know how they

evolved. Nevertheless, while the source constructions of grammatical markers are frequently unknown, there is enough evidence in our database (and the historical literature) to propose some plausible scenarios of constructional change for most of the demonstrative clause linkers in our sample.

In what follows, we analyse eight different types of clause-linking morphemes that are frequently derived from a demonstrative. We begin with relative pronouns (§2), which have been very prominent in the older literature on grammaticalisation (Brugmann 1904; Bühler 1934: 402), and then turn to a wide range of other markers, including linking and nominalising articles (§3), quotative markers (§4), complementisers (§5), conjunctive adverbs (§6), adverbial subordinate conjunctions (§7), correlatives (§8), and topic markers (§9).

2 Relative pronouns

The term “relative pronoun” is used in various ways by different scholars (see Lehmann (1984: 248–252) and van der Auwera (1985) for discussion); but for the purpose of this study, we adopt the following definition: A relative pronoun is an anaphoric pronoun that represents the head noun at the beginning of a postnominal relative clause.

Relative pronouns are frequent in European languages, but rare outside of Europe (Comrie 2006). In our sample, there are four European languages in which relative pronouns descended from a question word (French, Georgian, Hungarian, Serbo-Croatian), and one language (German) in which relative pronouns are based on a demonstrative.¹ As can be seen in (1a)–(1c), German relative clauses are introduced by a demonstrative relative pronoun that indicates the syntactic function of the head noun through case-marking or a preposed adposition.

- (1) Modern German (Indo-European, Germanic)
 - a. Das ist der Mann, **der** mir geholfen hat.
this is the man REL.NOM me helped has
'This is the man who helped me.'
 - b. Das ist der Mann, **den** ich gesehen habe.
this is the man REL.ACC I saw have
'This is the man who I saw.'

¹In Hungarian, relative pronouns are derived from question words by the prefix *a-*, which is historically related to the demonstrative *az* ‘that’, e.g. *a-ki* ‘THAT-who’ (Kenesei et al. 1998: 40).

- c. Das ist der Mann, mit dem ich gesprochen habe.
 this is the man with REL.DAT I spoken have
 'This is the man who I talked to.'

Outside of Europe, there are only four other languages in our sample in which relative pronouns are introduced by a demonstrative that qualifies as a relative pronoun according to our definition. One of them is Tümpisa Shoshone (2).

(2) Tümpisa Shoshone (Uto-Aztecán, Numic; Dayley 1989: 358, 359)

- a. Wa'ippü nia pusikwa [atü hupiatüki-tü].
 woman.SBJ me know that.SBJ sing-PTCP.PRS.SS.SBJ
 'The woman who is singing knows me.'
- b. Wa'ippüa nüü pusikwa [akka hupiatüki-tünna].
 woman.OBJ 1SG know that.OBJ sing-PTCP.PRS.SS.OBJ
 'I know the woman who's singing.'

In Tümpisa Shoshone, relative clauses are commonly introduced by a case-marked demonstrative pronoun which Dayley (1989: 357) classifies as a "relative pronoun". Note, however, that while the Shoshone relative pronouns are inflected for case (and number), like those in many European languages, they do not signal the syntactic function of the head within the relative clause but agree in case (and number) with the preceding noun.

Another language in which relative clauses are introduced by demonstratives that may be analysed as relative pronouns is Yagua, an Amazonian language of Peru. There are two relative markers in Yagua (3) (Payne & Payne 1990: 342–346): (i) a "relative particle" that consists of the demonstrative *jirya* and the second position clitic *-tiy*, and (ii) a set of "relative pronouns" that agree with the preceding head in class and number. Note that the "relative pronouns" are not inflected for case, but in oblique relatives, demonstratives (or third person pronouns) are combined with bound adpositions that specify the syntactic role of the head in the relative clause (3b), like oblique relative pronouns in German (1c).

(3) Yagua (Peba-Yaguan; Payne & Payne 1990: 345, 346)

- a. vánú [jiy-ra-tiy ray-díiy-tániy-jáy jantya-sjj-níí] ...
 man this-CLF.N-REL 1SG-see-CAUSE-PROX imitate-NMLZ-3SG
 'The man I showed the picture ...'
- b. sa-ráváqá [rá-mu-tiy riy-púýtya-jada jáyánú-miy] ...
 3SG-poison INAN-LOC-REL 3PL-paint-PST fer.de.lance-PL
 'His poison in which the fer-de-lances painted themselves ...'

Similar types of relative pronouns occur in Tamashek, a Berber language of Mali and Algeria, in which relative clauses are introduced by a demonstrative that hosts an adposition clitic if the head serves an oblique role in the relative clause (4a)–(4b).

- (4) Tamashek (Afro-Asiatic, Berber; Heath 2005: 633, 636)
- a. é-hæn [w-á=dæx t-àzəbbu-y]
M.SG-house M-DEM.SG=in IMPF-go.down.IMPF-1SG.SBJ
'The house in which I go down (= spend the night).'
 - b. æ-háləs [w-á=s Ø-æbà rure-s].
M.SG-man M-DEM.SG=INS 3M.SG.SBJ-be.lost.PFV son-3SG.POSS
'The man whose son was lost (= died).'

Since postnominal relative clauses including a relative pronoun are similar to paratactic sentences, it is often assumed that relative pronouns are derived from anaphoric demonstrative pronouns of structurally independent sentences that have been downgraded to subordinate clauses (Heine & Kuteva 2007: 224–229; Givón 2009: 105). The hypothesis is not implausible, but difficult to verify by concrete diachronic data (Harris & Campbell 1995: 282–286). In fact, the diachronic data suggest that relative clauses typically develop under the influence of multiple source constructions (Hendery 2013). For instance, Lockwood (1968) argued that the relative clauses of Modern German are related to an old *apo koinou* construction in which a demonstrative pronoun served a double role in main and subordinate clauses (5) (see also Pittner 1995).

- (5) Old High German (Indo-European, Germanic; Lockwood 1968: 243)
- | | | | | | | |
|------|--------|---------|---------|-----|----------|--------|
| thô | liefun | sâr | thie | nan | minnôtun | meist. |
| then | ran | at.once | DEM.NOM | him | loved | most |
- 'Then ran at once those who loved him most.'

The sentence in (5) includes a demonstrative that serves as subject of two verbs: *liefun* 'ran' and *minnôtun* 'loved'. According to Lockwood (1968: 242–244), *apo koinou* constructions are easily extended to relative clauses if the two roles of the demonstrative are expressed by separate (pro)nouns (cf. *Wer ist die, die aufgeht aus der Wüste* 'Who is the one who rises from the desert'; see also Paul (1916–1920: IV: 189–191)). Since constructions of this type were frequent in Old and Middle High German, it is not implausible that they influenced the development of relative clauses; but that does not mean that paratactic sentences did not also impact their development. As Hendery (2013) has shown, relative clauses are

often historically related to more than one source. In the current case, we know that relative pronouns often develop from demonstratives, but this development may involve demonstratives in several source constructions (cf. Lehmann 1984: 378–383; Diessel 1999a: 120–123).

3 Linking and nominalising articles

Since subordinate clauses are frequently expressed by nominalisations (Lehmann 1988), they are often marked by the same morphemes as noun phrases. For instance, in many languages subordinate clauses are accompanied by articles or determiners that one might analyse as particular types of clause linkers. Dryer (1989) defined the term “article” by two features: (i) articles are used to indicate (in)definiteness and/or (ii) serve as formal markers of noun phrases. The articles of subordinate clauses are of the latter type. They are formal markers of nominal constituents but do not indicate (in)definiteness.

Two basic types of subordinating articles may be distinguished: (i) linking articles and (ii) nominalising articles. The two types of articles form a continuum, but for the purpose of this study we reserve the term “linking article” for markers that are primarily used to combine a head noun with attributes, and we use the term “nominalising article” for markers that are primarily used to form nominal constituents. Crucially, both types of articles are commonly derived from demonstratives. In many Austronesian languages, for example, attributes are linked to the head noun by an article, as in (6) from Toba Batak.

- (6) Toba Batak (Austronesian, Malayo-Polynesian, Foley 1980: 186)
- a. bijang **na** balga
dog LK big
'a big dog'
 - b. baoa **na** mang-arang buju i
man LK ACT-write book the
'the man who wrote the book'

As can be seen, the adjective in (6a) and the relative clause in (6b) are linked to a preceding noun by the marker *na*, which Foley (1980: 186–187) calls a “ligature” and Himmelmann (1997: 173) a “linker” or “linking article”. Similar types of linking articles occur in many other Austronesian languages, including Tagalog, Wolio and Ilokano. In all of these languages, relative clauses are linked to the preceding noun by the marker *na* or *a*, which is historically related to the medial

demonstrative **a/na* of Proto-Austronesian (Himmelmann 1997: 164; Ross 1988: 100). While *a/na* is also used with adjectives and other types of noun modifiers, it is particularly frequent with relative clauses (Foley 1980).

Linking articles are very common in the Austronesian language family, but are also found in many other languages across the world. Schuh (1983; 1990) and Hetzron (1995) showed that they are widely used in Chadic, Cushitic and Semitic languages, and Aristar (1991) presented data from a wide range of languages in which relative clauses and genitive attributes are marked by the same linker. All of these studies emphasise that linking articles are very frequent with relative clauses and commonly derived from demonstratives.

Like linking articles, nominalising articles are often based on demonstratives. Consider, for instance, the examples in (7a)–(7c) from Chumash, in which relative clauses are syntactic nominalisations marked by the article *l=* and the dependent proclitic *hi=*, which, according to Wash (1999: 46), is based on a demonstrative. Since nominalised clauses serve as syntactic NPs, they can be used without a nominal head as free relatives (7c).

- (7) Chumash (Isolate; Wash 2001: 76, 97, 77)

- a. *hi=l=xip*
DP=ART=rock
'a/the rock'
- b. *hi=l=xip-xip-?* *hi=l=?-iy-sa?-išmax-šiš*
DP=ART-rock-rock-EM DP=ART=NMLZ-PL-FUT-throw.at-RECP
'(and) rocks that they can throw at one another'
- c. *?i=s-uškál* *hi=l=?-iy-qili?-aqmil.*
TOP=3-be.strong DP-ART-NMLZ-PL-HAB-EP-drink
'What they used to drink was strong.'

Similar types of nominalising articles (derived from demonstratives) occur in other languages of our sample. In Jamul Tiipay, for instance, nominal clauses and internally-headed relatives are marked by the demonstrative clitic =*pu* (8b)–(8c), which also occurs with nouns (8a). Note that the demonstrative clitic is a determiner that cannot be used as an independent pronoun like the demonstratives of many other languages (Diessel 2005a), and that =*pu* is followed by a case clitic if the subordinate clause functions as subject of the main verb (8c).²

²Jamul Tiipay is a “marked nominative language” in which subjects are marked by a case morpheme, whereas objects are “zero-marked” (Comrie 2013).

- (8) Jamul Tiipay (Hokan, Yukan; Miller 2001: 153, 220, 208)

- a. **wa=pu**
house=DEM
'that house'
- b. [puu-ch wi'i-x]=pu uuyaaw.
that.one-SBJ do-IRR=DEM know
'I know she will do it.'
- c. ['iipa peya nye-kwe-'iny]=pu=ch mespa.
man this 3/1-SJREL-give=DEM=SBJ die
'The man who gave me this died.'

Very similar types of nominal and internally headed relative clauses occur in other languages of our sample. In Assiniboine, for example, nominalised subordinate clauses are marked by the distal demonstrative *žé* 'that' (9), or, less frequently, by the proximal demonstrative *né* 'this' (Cumberland 2005: 415–417).

- (9) Assiniboine (Siouan; Cumberland 2005: 347, 415, 417)

- a. [wíya žé] Ø-háska.
woman that A3-be.tall
'That woman (over there) is tall.'
- b. John [mnatkj-kte-ši žé] snok-Ø-yá.
John Ø-A1SG-drink-POT-NEG that ST-A3-know
'John knows that I'm not going to drink it.'
- c. [wjč^hášta t^himáni Ø-hí žé] mi-nékši Ø-é.
man visit A3-arrive.here that 1.POSS-uncle A3.be
'The man who came to visit us is my uncle.'

According to Schuh (1990) and Aristar (1991), linking articles are often based on demonstrative pronouns that were originally used as heads of complex NPs. The best evidence for this development comes from Akkadian, an old Semitic language of Mesopotamia with extensive diachronic records (Deutscher 2000; 2009).

Like many other Afro-Asiatic languages, Akkadian had a linking article that occurred with nominal attributes. In Old and Middle Babylonian (1950 BC to 1000 BC), the linker *ža* was an invariable marker, but this marker developed from the demonstrative pronoun *šu*, which was inflected for gender, number and case. Analysing data from Old Akkadian (2500 BC to 1950 BC), Deutscher

(2001; 2009) showed that *šu* was originally the pronominal head of a genitive attribute that was later extended to relative clauses. Both genitive attributes and relative clauses were frequently used with a demonstrative pronoun as head in Old Akkadian but, crucially, in the course of the development, *šu* lost its status as a pronoun and turned into a formal marker of certain types of attributes. Since *šu* was originally the head of a complex NP, the new genitive and relative constructions marked by *šu* (or *ža*) could be used without a co-occurring noun as syntactic nominalisations. Nevertheless, since the *šu*-nominalisations were often used in apposition to a preceding noun, they regained their original function as noun modifiers (10).

- (10) $[[\dot{sh}u]_{PRN} [GEN \text{ or } RC]_{MOD}]_{NP} > [\dot{sh}u \text{ GEN or RC}]_{NP} > [NP]_{NP} [\dot{sh}u \text{ GEN or RC}]]_{NP}$

The development of the Akkadian linker provides a plausible account for many of the properties that are characteristic of linking and nominalising articles: It explains why relative clauses are often marked by the same demonstrative linker as genitive attributes and why in many languages relative clauses can be used without a (pro)nominal head as free nominals or syntactic nominalisations (Schuh 1983; 1990; Aristar 1991).

4 Quotative marker

A quotative marker is defined as a conjunction-like element that serves to mark direct speech. In some languages, quotative markers are based on general speech verbs meaning ‘say’, ‘talk’ or ‘speak’. The development of quotative markers from speech verbs has been very prominent in early research on grammaticalisation (Lord 1993; Klamer 2000); but, as Güldemann (2008) showed, based on data from African languages, quotative markers are also frequently derived from manner demonstratives. This is confirmed cross-linguistically by our data.

Manner demonstratives are a particular subclass of demonstratives that serve to draw interlocutors’ attention onto the manner of an action (König 2012; König this volume; Nikitina & Treis this volume; Teptiuk this volume). In English, manner demonstratives are complex forms consisting of the similitative marker *like* and a demonstrative pronoun (e.g. *He did it like this*); but in many other languages, manner demonstratives are simple lexemes, which may or may not be formally related to demonstrative pronouns. In German, for instance, the manner demonstrative *so* is formally distinct from demonstrative pronouns, but in Ambulas, a Sepik language of Papua New Guinea, manner demonstratives include the same deictic roots as all other demonstratives (see Table 1).

Table 1: Demonstratives in Ambulas (Wilson 1980: 56–57)

	Pronouns	Determiners	Locative	Manner
Proximal	<i>dé-kén</i> ‘3SG-this’	<i>kéni</i> ‘this’	<i>kéba</i> ‘here’	<i>kéba</i> ‘so/thus’
Distal	<i>dé-wan</i> ‘3SG-that’	<i>wani</i> ‘that’	<i>waba</i> ‘there’	<i>waga</i> ‘so/thus’

Like all other demonstratives, manner demonstratives can refer to entities in the surrounding speech situation, but there seems to be a general tendency to use them with reference to sentences or propositions (König 2012). In particular, manner demonstratives are often used to indicate direct speech, as in (11) and (12) from German and French.

- (11) German (Indo-European, Germanic)

Ja, ich würde das **so** sagen: “Das ist ein Sonderfall.”
 yes I would that so say this is a special.case
 ‘Well, I would put it this way: “This is a special case.”’

- (12) French (Indo-European, Romance)

Marie s'est exprimée ainsi: “Puisqu'il le faut, j'irai.”
 Marie REFL.is express.PTCP thus since.it.must.be I.go.FUT
 ‘Marie expressed herself in this way: “Since it is necessary, I will go.”’

Similar uses occur in many other languages of our sample. For instance, in Bariai, an Austronesian language of New Britain, the verb *keo* ‘say’ is frequently accompanied by a manner demonstrative to mark direct speech (13).

- (13) Bariai (Austronesian, Malayo-Polynesian; Gallagher & Baehr 2005: 157)

Taine toa oa i-keo pa-n **bedane**, “Gergeu ne taine”.
 female GIVEN there SBJ.3SG-say at-3SG.OBL like.this child here female
 ‘That woman spoke to him like this, “This child is a girl.”’

Interestingly, some languages use different types of manner demonstratives for previous and subsequent quotations (Teptiuk this volume). In Ambulas, for example, *kéba* ‘so/thus’ refers to a subsequent quote, whereas *waga* ‘so/thus’ refers backwards. A parallel contrast occurs in Usan, a Papuan language of New Guinea, in which *ete* ‘thus’ is used to announce upcoming speech (14), whereas *ende* ‘so/thus’ refers to a preceding quotation (see also Korafe; Farr 1999: 276).

- (14) Usan (Trans-New Guinea, Madang; Reesink 1984: 184)
- munon eng ete yo-nob qâm-ar: “mâni âib ne-teib-âm,”
 man the thus me-with say-3SG.PST food big you-give.SG.FUT-1SG
ende qâm-arei.
 thus say-3SG.PST
- ‘The man said thus to me: “I will give you a lot of food,” thus he said.’

There is a fluid transition between the discourse-deictic use of manner demonstratives and grammaticalised quotative markers. In the examples considered thus far, the demonstratives are only weakly grammaticalised. Yet, there are languages in which manner demonstratives have developed into true quotative markers. Meithei, for example, has “quotative complementizers” (Chelliah 1997: 190) that are derived from the verb *háy* ‘say’, the nominaliser *-pə* and a demonstrative clitic, i.e. *=si* PROX or *=tu* DIST (15).

- (15) Meithei (Sino-Tibetan, Kuki-Chin; Chelliah 1997: 305)
- Tomba-nə Tombi-nə má-pu ñay-həw-li **háy-pə=du**
 Tomba-CT Tombi-CT he-PAT wait-START-PROG say-NMLZ=that
 kaw-thok-ləm-í.
 forget-OUT-EVD-NHYP
- ‘Tomba forgot that Tombi had been waiting for him.’

Like Meithei, Thai Kamti has grammaticalised quotative markers that are derived from the verb *waa*³ ‘say’ and the proximal demonstrative *nai*¹ ‘this’. The two morphemes have fused into one word that is often reduced to *wan*¹ in quotative constructions (16).

- (16) Thai Kamti (Tai-Kadai, Kam-Tai; Inglis 2014: 123)
- “maeu⁴ mai³ man⁴ khaeu³ han⁵ uu⁵” **wan**¹.
 2SG OBJ 3SG want see IMPF QUOTE
- ‘He says that he wants to see you.’

Note that the quotative marker in (16) is not accompanied by a speech verb. Since *wan*¹ includes the verb *waa*³ ‘say’, one might think of *wan*¹ as some kind of verb, but it is not unusual that quotative markers are used without a verb. In German, for example, the manner demonstrative *so* can refer to direct speech without a co-occurring verb (17) (Golato 2000).

- (17) German (Indo-European, Germanic)
 Und ich **so**: “Okay, das ist deine Chance.”
 and I thus okay this is your chance
 ‘And I am like, “Okay, this is your chance.”’

Similar types of non-verbal quotative clauses occur in other languages of our sample, as for instance in Komnzo (18), a Yam language of Papua New Guinea.

- (18) Komnzo (Yam, Tonda; Döhler 2018: 331)
 naf **nima** “Nakre, wimäs=en mni
 3SG.ERG like.this Nakre mango.tree=LOC fire
 b=ŋasog.”
 MED=2:3.SBJ.NPST.IPFV.climb
 ‘He (said): “Nakre! The fire is climbing up the mango tree.”’

Interestingly, Güldemann (2008: 322–326) argued that demonstrative quotative markers can acquire properties of verbs when they are routinely used in non-verbal clauses to mark direct speech. In Epena Pedee, for instance, the manner demonstrative *má-ga* ‘that-like’ may be inflected for tense if it is not accompanied by a speech verb (19).

- (19) Epena Pedee (Choco; Harms 1994: 63, 176)
- a. **má-ga** hara-hí, “...”
 that-like tell-PST
 ‘He said as follows ...’
 - b. **má-ga-hí**, “pháta kho-páde a-hí.”
 that-like-PST plantain eat-IMP say-PST
 ‘That is: “Eat your plantains.”’

The data from Epena Pedee provide good evidence for Güldemann’s (2008: 529) claim that quotative constructions provide “a highly fruitful cradle of new verbs”.

5 Complementisers

In formal syntax, a complementiser is a particular word class category that serves as head of a “complementizer phrase” (Radford 1997). However, in what follows, we use the term “complementiser” in a more traditional way for subordinate

conjunctions of nominal clauses functioning as subject or object of the main verb.

Like many other types of clause linkers, complementisers are often based on demonstratives. English *that* and German *dass* are well-known examples. There are several other languages in our sample in which nominal clauses are marked by a demonstrative. In fact, we have already seen some examples in §3. Recall that the nominal clauses of Jamul Tiipay and Assiniboine are marked by a clause-final demonstrative (8).

The position of the complementiser correlates with the order of verb and object and the position of the nominal clause relative to the main verb (Schmidtke-Bode & Diessel 2017). In OV languages, nominal clauses usually precede the main verb and include a clause-final complementiser, as in Jamul Tiipay and Assiniboine, whereas in VO languages, nominal clauses typically follow the main verb and are marked by an initial complementiser, as in English and German. There are several other languages with initial and final demonstrative complementisers in our sample. Consider, for instance, (20) and (21) from Amele and Tamashuk.

- (20) Amele (Trans-New Guinea, Madang; Roberts 1987: 47)
[Naus uqa uqa na ho qo-i-a eu] ija d-ug-a.
Naus 3SG 3SG of pig hit-3SG-PST that 1SG know-1SG-PST
'I know that Naus killed his pig.'
- (21) Tamashuk (Afro-Asiatic, Berber; Heath 2005: 674)
ànne-γ=a-s [à=d i-nzər]
say.PFV-1SG.SBJ=DAT-3SG DEM=COM 3M.SG.SBJ-sing.IMPF
'I told him to sing.' (Lit. 'I said to him, that he sing.')

Amele is an OV language in which nominal clauses precede the main clause predicate, and Tamashuk is a VO language in which nominal clauses are postposed to the main verb. As can be seen, like Jamul Tiipay and Assiniboine, Amele marks preverbal nominal clauses by a clause-final demonstrative; like English and German, Tamashuk marks postverbal nominal clauses by a clause-initial demonstrative. Other languages in which complementisers are based on demonstratives include Chumash, Lakhota and Diegueño.

Note that while demonstrative complementisers are not uncommon, they are less frequent than many other types of demonstrative clause linkers in our database. In particular, the markers of relative clauses are more often based on demonstratives than the markers of nominal clauses. Concentrating on those

markers for which we were able to determine a diachronic source, more than 50% of all (free) relative markers are based on demonstratives in our data, whereas only about 15% of all complementisers are related to demonstratives. What is more, with one exception (see below), all of the demonstrative complementisers included in our database also occur in relative clauses, suggesting that complementisers and relativisers are historically related (e.g. English *that*).

In the literature it is often said that demonstrative complementisers derive from discourse-deictic demonstratives (Harris & Campbell 1995: 287). In particular, it is widely assumed that the German complementiser *dass* developed from a paratactic demonstrative pronoun (Behaghel 1928: 30; Ebert 1978: 26). According to the standard analysis, *dass* has evolved from a cataphoric demonstrative that served to anticipate an upcoming sentence as in *Listen to this: John and Sue will get married*. On this account, the grammaticalisation of *dass* involved several related changes whereby a cataphoric demonstrative pronoun turned into a formal marker of the subsequent sentence that was downgraded to a subordinate clause. This analysis is based on the occurrence of the demonstrative *thaz* in two different structural positions in Middle High German (22).

- (22) Middle High German (Indo-European, Germanic; Axel 2009: 25, 26)

- a. Joh gizálta in sar **tház** \ thiu sálida untar ín was.
and told them immediately that he luck among them was
'And told them immediately that good fortune was among them.'
- b. "Íh," quad er, "infúalta \ **thaz** étheswer mih rúarta;" ...
I said he felt that someone me touched
"I," he said, "feel, that someone touched me;" ...

In (22a) *thaz* occurs at the end of the first sentence and seems to anticipate the subsequent clause, and in (22b) *thaz* occurs at the beginning of the second sentence where it seems to serve as a formal marker of a nominal clause. Given that some authors of that period used the demonstrative *thaz* in both ways (e.g. Otfrid), it seems plausible to assume that the alternation between the two uses of *thaz* reflects ongoing syntactic change.

However, several recent studies have questioned this view (Lühr 2008; Axel 2009; Schmidtke-Bode 2014). According to Axel (2009), there is little evidence for the cataphoric use of demonstrative pronouns in Middle High German. The few examples that are commonly cited to illustrate this use, notably (22a), are unclear and leave room for alternative interpretations (Axel 2009: 25). Challenging the traditional view, Axel and Lühr suggest that the complementiser *dass* did

not develop from a cataphoric demonstrative but from a relative pronoun. In particular, they argue that *dass* emerged in the context of a correlative construction in which the relative pronoun *thaz* occurred together with a demonstrative or correlative pronoun in the preceding main clause (23).

- (23) Middle High German (Indo-European, Germanic; Axel 2009: 29)
- Er tháhta odowila **tház** \ **thaz** er ther dúriwart wás.
he thought maybe that that he the doorkeeper was
'He thought that maybe he was the doorkeeper.'

Correlative constructions of this type were frequent in Middle High German and provide a plausible bridging context between relative and nominal clauses. Moreover, the scenario that Axel and Lühr suggest for German is consistent with the scenario that has been proposed for other languages in which relative and nominal clauses include the same marker (Givón 1991; Schmidtke-Bode 2014: 248–254). As pointed out above, if nominal clauses are marked by a demonstrative, relative clauses often include the same demonstrative, which is readily explained if we assume that demonstrative complementisers derive from demonstrative relativisers.

Nevertheless, there is a second scenario whereby a demonstrative pronoun may develop into a complementiser. As Lord (1993) and others have shown, complementisers are frequently derived from quotative markers. Since quotative markers are often based on manner demonstratives (cf. §4), it is a plausible hypothesis that complementisers may develop from demonstratives via quotative constructions. The grammaticalisation literature has concentrated on the development of complementisers from speech verbs, but there is at least one language in our sample in which a complementiser (that does not also occur in relative clauses) may have evolved from a demonstrative quotative marker. In Noon, direct and indirect speech are marked by the "manner adverb" *an* meaning 'thus' or 'in this way' (24a). Since *an* is also used as a complementiser with verbs of cognition (24b)–(24c), it is not unreasonable to assume that the complementiser use of *an* has developed from its use in quotative constructions.

- (24) Noon (Niger-Congo, Atlantic; Soukka 2000: 314)
- a. Yaal-aa hay-ya, wo'-a-ri **an:** "Mi hot-in ee-fu."
man-DEF come-NARR say-NARR-OBJ.3SG thus I see-PFV mother-2SG
'The man came and said to him/her (this): "I've seen your mother."

- b. Ya halaat-ee an: “Mi hot-oo ken.”
s/he think-PST COMP I see-PRES.NEG nobody
‘S/he thought (this): “I don’t see anybody.”’
- c. Cica foog-ee an baa keloh-hii-ri.
grandmother think-PST COMP individual hear-ASP.NEG-OBJ.3SG
‘Grandmother thought that the person hadn’t heard her.’

In general, complementisers are historically related to demonstratives, but it seems that this relationship is usually mediated by the use of demonstratives in relative and quotative constructions. In particular, the extension of demonstrative relative markers to demonstrative complementisers is cross-linguistically very common (Schmidtke-Bode 2014: 248–254).

6 Conjunctive adverbs

Conjunctive adverbs are paratactic clause linkers that combine two independent sentences. In contrast to many other types of clause linkers, they have received little attention in typology. In studies of English grammar, the term “conjunctive adverb” applies to discourse connectives such as *however*, *thus* and *nevertheless*. Similar types of discourse connectives occur in many other languages and often involve demonstratives. In what follows, we provide an overview of the conjunctive adverbs in our database concentrating on those forms that involve demonstratives. As we will see, conjunctive adverbs vary along several dimensions:

1. They can be more or less complex ranging from mono-morphemic words to (frozen) multi-word expressions.
2. They are usually associated with the second conjunct but exhibit different degrees of formal integration.
3. They express a wide range of semantic relations including, above all, relations of time, cause and reason.

In some languages, conjunctive adverbs are based on manner demonstratives. In English, for example, the manner demonstratives *so* and *thus* are commonly used as conjunctive adverbs that designate a consequence or logical conclusion (25). Likewise, Finnish *niin* ‘so/thus’ and Japanese *koo/so/aa* ‘in this/that way’ are manner demonstratives that can be used as conjunctive adverbs (König 2012).

- (25) He failed the exam; **thus/so**, he had to repeat the class.

Apart from manner demonstratives, oblique demonstrative pronouns provide a common source for conjunctive adverbs. In Yurakaré, for example, temporal clauses are introduced by *latijsha*, which is composed of three morphemes: the endophoric reference marker *l-*, the anaphoric medial demonstrative *ati* and the ablative case marker *=jsha* (26).

- (26) Yurakaré (Isolate; van Gijn 2006: 321)

mi-bëjtì	së=ja	latijsha	shuyuj-ta-m.
2SG-see.1SG.SBJ	1SG.PRN=EMPH	then	hidden-MID-2SG.SBJ
'I saw you, then you hid yourself.'			

Santali also uses oblique demonstratives to indicate sequential links between two structurally independent sentences. Result and causal clauses are introduced by *ente* 'because/for' or *onate* 'therefore/so.that', which are based on the demonstratives *en* 'that' and *ona* 'that.INAN' and the instrumental suffix *-te* (27).

- (27) Santali (Austro-Asiatic, Munda; Neukom 2001: 180)

am-then-ge	baba-n	cal-a-k'-kan-a;	[ɛnte=j]
2SG-DAT-FOC	father-1SG.SBJ	go-MID-IPFV-IND	that.INS=1SG.SBJ
badae-y-et'-a ...].			
know-EP-IPFV-IND			

'I am coming to you, father, because I know ...'

Functionally equivalent to oblique demonstratives are adpositional constructions consisting of a demonstrative pronoun or adverb and an adposition. English *therefore*, for instance, derives from Old English *pærfore* 'for that', which is composed of the demonstrative *pær* 'there' and the adposition *fore* 'before, because of'. Similar types of conjunctive adverbs occur in many other languages of our sample. Some examples are given in Table 2.

Conjunctive adverbs of this type are commonly used to indicate relations of cause, reason and time. Some of these expressions may still be seen as adpositional constructions, but others have turned into monomorphemic clause linkers. At the initial stage of the development, the demonstrative directs the interlocutors' attention to a previous clause or proposition and the adposition specifies a particular semantic relationship between the two clauses. Yet, as the development continues, the demonstrative and the adposition may lose their status as independent words and may fuse into a single morpheme (e.g. German *darum*).

There are also some languages in our sample in which conjunctive adverbs are based on demonstratives and topic or focus markers. In Galo (Sino-Tibetan),

Table 2: Examples of conjunctive adverbs

Language	Form	Gloss	Translation
German	<i>darum</i> (< <i>da:r-umbi</i>)	that.OBL-because.of	‘therefore’
Japanese	<i>sore-kara</i>	that-after	‘and then’
Burmese	<i>da=jaun</i>	this=because.of	‘therefore’
Awa Pit	<i>suna=akwa</i>	that=because.of	‘because of that’
Supyire	<i>lire e</i>	this in	‘so, therefore’
Epena Pedee	<i>maa-p^héda</i>	like.that-after	‘after that’
Menya	<i>i-ta-ŋi</i>	that-from-given	‘after that, as a result’
Hixkaryana	<i>ire ke</i>	that because.of	‘therefore’
Koyra Chiini	<i>woo di banda</i>	DEM DEF behind	‘afterwards’
Chumash	<i>?akim-pi</i>	there-LOC	‘during (that time)’

for instance, sequential relations of time and result are expressed by *okkəə* ‘and, then, so’, which derives from the ablative demonstrative *okə* ‘this.NEAR.YOU’ and the topic marker *əə* (Post 2007: 370). Similarly, in Bilua some “linking adverbs” are based on a distal demonstrative and a focus marker (28).

- (28) Bilua (Solomons East Papuan; Obata 2003: 45)
- a. *sainio* ‘therefore, then’ < *sai inio* ‘there FOC’
 - b. *sainio* ‘therefore, accordingly’ < *so inio* ‘that FOC’

Apart from manner demonstratives and adpositional phrases, linking clauses provide a common source for conjunctive adverbs. There are various types of linking clauses (cf. Guérin 2019), but many of them are organised around a demonstrative and a proverb verb such as ‘be’ or ‘do’, as in (29) and (30) from Alamblak and Manambu.

- (29) Alamblak (Sepik, Sepik Hill; Bruce 1984: 283)
- | | |
|---------------------------------|------------------|
| yira buga-m fa-më-r-m. | [ind-net-r-n, |
| fish all-3PL eat-RPST-3SG.M-3PL | DEM-do-3SG.M-DEP |
| yati-fa-më-r.] | |
| stomach-eat-RPST-3SG.M | |
- ‘He ate all the fish. He did that (therefore), he had a stomach ache.’

- (30) Manambu (Sepik, Middle Sepik; Aikhenvald 2008: 494)

sana:k karabə ja:p kui-taka-dana-ti,
money.LK.DAT men's.house.LK thing give.to.third.PST-put-3PL.SBJ-3PL
[alək tə-ku ma: tə].
that.DAT BE-ASP.SS NEG HAVE.NEG

'They gave away the things from men's house for money, this is why they do not have (them anymore)'.

In both examples, the second sentence is connected to the preceding sentence by a (linking) clause that includes a demonstrative pronoun and a proverb. In Alamblak, the linking clause is a single word consisting of the demonstrative *ind*, the proverb *net* 'do', a third person suffix and a dependent marker. In Manambu, the linking clause is composed of the distal demonstrative *a* 'that' (in dative case) and a medial clause including the verb *tə* 'be, stand'.³ Note that both the demonstrative and the medial clause are also used alone for clause combining, but according to Aikhenvald (2008: 494) the expression *alək tə-ku* is in the process of developing into a complex clause linker meaning 'and so, as a result'.

Linking clauses of this type provide a common strategy of clause combining in Papuan languages (e.g. Alamblak, Manambu, Korafe, Menya) but also occur in other languages in our sample. Korean, for instance, has a whole series of "conjunctive adverbials" that derive from linking clauses including the demonstrative *ku* 'that' and the verb *ha(y)* 'do, be' (31).

- (31) Korean (Isolate; Sohn 1994: 89–90; 2009: 292)

- a. *kulayse* 'so, thus, therefore' < ***ku-li/le-hay-se***
'that-along/like-do/be-as'
- b. *kuliko* 'and' < ***ku-li-ha-ko*** 'that-along/like-do-and'
- c. *kulehciman* 'but, however' < ***ku-li/le-ha-ciman***
'that-along/like-do/be-though'
- d. *kulinikka* 'therefore' < ***ku-li/le-ha-nikka***
'that-along/like-do/be-because'

³Medial clauses are dependent clauses of clause chaining constructions. They occur with switch-reference markers that indicate whether the subsequent clause includes the same or a different subject (Haiman & Munro 1983).

7 Adverbial subordinate conjunctions

Adverbial clauses are subordinate clauses that express a wide range of semantic relations (Thompson et al. 2007; see also Diessel 2019b). Since many of these relations are also expressed by adpositional phrases, it is not surprising that adverbial clauses are often marked by adpositions. In English, for example, some temporal adverbial clauses are introduced by subordinate conjunctions that are also used as temporal prepositions (e.g. *since*, *after*, *before*).

Across languages, there is a close connection between adverbial subordinators and certain semantic types of adpositions, notably adpositions of time, cause and purpose. However, in addition to adpositions, adverbial clauses occur with a wide range of other subordinating morphemes, including morphemes that are historically related to demonstratives. In German, for example, some adverbial clauses of time and purpose are introduced by subordinate conjunctions that include the demonstratives *dem* ‘that.DAT’, *da* ‘there’ and *so* ‘so, thus’ (32).

- (32) Modern German (Indo-European, Germanic)
 - a. *seitdem* ‘since’, *nachdem* ‘after’, *indem* ‘by’
 - b. *damit* ‘in order to, so that’, *da* ‘since, as, because’
 - c. *sobald* ‘as soon as’, *sofern* ‘as long as’

Note that some of the subordinate conjunctions in (32) are composed of a demonstrative and an adposition, similar to conjunctive adverbs such as *darum* ‘therefore’ (see §6). We will come back to this below. Here we note that while subordinate conjunctions are often similar to conjunctive adverbs, there is a clear structural difference between them in Modern German. In contrast to conjunctive adverbs (e.g. *darum*), adverbial subordinators (e.g. *damit*) introduce subordinate clauses that are distinguished from main clauses, or paratactic sentences, by a particular word order. As can be seen in (33a), in adverbial clauses, the finite verb occurs in clause-final position, whereas in main clauses (33b), the finite verb comes in second position, i.e. right after the conjunctive adverb. Thus, while it is often said that adverbial clauses and paratactic sentences form a continuum (e.g. Thompson et al. 2007: 237), there are languages like German in which the continuum is divided into separate constructions.

- (33) Modern German (Indo-European, Germanic)
 - a. Wir gehen jetzt, **damit** wir nicht zu spät sind.
we go now so.that we not too late are
'We are leaving now so that we won't be late.'

- b. Wir haben den Zug verpasst; **darum** sind wir zu spät.
we have the train missed therefore are we too late
'We missed the train; that's why we are too late.'

Apart from German, there are several other languages in our sample in which adverbial clauses are marked by subordinate conjunctions that are etymologically related to demonstratives. For example, in English, result clauses are introduced by *so that*, and in French, conditional clauses are introduced by the manner demonstrative *si* 'if' (König 2012). Two further examples of demonstrative subordinate conjunctions are shown in (34) and (35).

- (34) Tamashek (Afro-Asiatic, Berber; Heath 2005: 663)
[dà=s Ø-æmmu-t] n-ègla.
DEM=INS 3M.SG.SBJ.die.ASP-AUG 1SG.SBJ-go.away.ASP
'When he died, we went away.'
- (35) Yimas (Lower Sepik-Ramu, Lower Sepik; Foley 1991: 453)
[m-n-awram-r-mp-n] mpa-n namarawt anak.
DEM-3SG.A-enter-PFV-SG-OBL one-SG person.SG COP.1SG
'When he went in, he went alone.'

As can be seen, in Tamashek temporal 'when' clauses are marked by a demonstrative and an instrumental case clitic (Heath 2005: 663), and in Yimas "finite oblique clauses", which are functionally equivalent to adverbial clauses in English, are expressed by nominalisations that begin with the "near distal deictic base" *m-* 'that' (Foley 1991: 435). Other languages in which some adverbial subordinate conjunctions are historically related to demonstratives include Wari (time clauses), Jamul Tiipay (time and purpose clauses) and Bilinarra (conditional clauses) (see Heine & Kuteva 2007: 250–251 for additional examples).

Given that the subordinate conjunctions of adverbial clauses are often similar to conjunctive adverbs (e.g. German *damit* 'so that' with *darum* 'therefore'), we may hypothesise that (some) subordinate conjunctions derive from paratactic clause linkers (Hopper & Traugott 2003: 185). However, while this hypothesis is not implausible, there is little evidence for it in our data. On the contrary, the available data suggest that the demonstratives of adverbial subordinators do not usually derive from paratactic clause linkers but from demonstratives of other types of subordinate clauses. In German, for instance, adverbial conjunctions such as *seitdem* 'since' and *nachdem* 'after' are not derived from conjunctive adverbs of paratactic sentences but from oblique relative clauses in Old and Middle

High German, e.g. *sît dem mâle daz* ‘since the time that’ (Lockwood 1968: 238). Similar types of adverbial subordinators occur in Tamashuk (Heath 2005) and Yimas (Foley 1991), in which adverbial clauses are marked by demonstratives that also occur in relative clauses. While there are no diachronic data to investigate the diachronic origins of adverbial subordinators in Tamashuk and Yimas, Heath (2005: 663–675) and Foley (1991: 435–444) make it clear that the adverbial clauses of these languages are derived from oblique relatives.

More research is needed to determine the diachronic trajectories of demonstrative subordinate conjunctions in adverbial clauses, but judging from the evidence in our database we suspect that the demonstratives of adverbial subordinate conjunctions are more frequently derived from demonstrative relativisers, complementisers or linking and nominalising articles than from discourse deictic demonstratives or paratactic clause linkers.

8 Correlatives

The notion of correlative is used in many different ways in linguistics (Lipták 2009). In the current study, we use the term “correlative” for pronominal and conjunctive elements of main clauses that serve to indicate the occurrence of an associated subordinate clause (or a particular element within the subordinate clause). Since subordinate clauses are commonly marked by a subordinating morpheme – a relativiser, complementiser or adverbial conjunction – a correlative is often used together with a subordinate marker. In conditional sentences, for instance, subordinate conjunctions are often paired with a correlative in the main clause (e.g. English *if/then*).

What is important in the context of the current chapter is that correlatives are very often based on demonstratives. Consider, for instance, the two following examples of conditional/concessive sentences from German (36) and Hungarian (37).

- (36) Modern German (Indo-European, Germanic)
- Auch wenn noch vieles unklar ist, (so) müssen wir doch jetzt handeln.
 Even though still much unclear is so must we still now act
 ‘Even though much is still unclear, we must act now.’

- (37) Hungarian (Uralic, Ugric; Kenesei et al. 1998: 51)
Ha Péter el-alszik, (akkor) Anna meg-haragszik.
if Peter PRE-sleeps then Anna PRE-is.angry
'If Peter falls asleep, Anna will get angry.'

In both languages the main clauses of conditional/concessive sentences are optionally introduced by a correlative. The German correlative *so* is a manner demonstrative that is also used in many other contexts (cf. §4 to §6),⁴ and the Hungarian correlative *akkor* 'then' is composed of the demonstrative *az* 'that' and the temporal suffix *-kor* 'at (the time)'.

Correlatives are not only used with conditional/concessive clauses; they also occur with other semantic types of adverbial clauses. The following examples from German, Hungarian and Georgian include demonstrative correlatives that serve to anticipate upcoming adverbial clauses of manner (38), cause (39) and result (40).

- (38) Modern German (Indo-European, Germanic)
Ich mache das **so**, wie du gesagt hast.
I do that so as you said have
'I will do it (in the way) as you said.'
- (39) Hungarian (Uralic, Ugric; Kenesei et al. 1998: 51)
Anna az-ért haragszik, mert Péter elaludt.
Anna that-for is.angry because Peter slept
'Anna is angry because Peter has fallen asleep.'
- (40) Georgian (Kartvelian; Hewitt 1995: 578)
Ik iset-i mgl-eb-i da t'ur-eb-i ar-i-an, rom
there like.that-AGR wolf-PL-NOM and jackal-PL-NOM be-PRS-they SUB
še-g-č'am-en.
PRE-you-devour-they.FUT
'There are such wolves and jackals there that they will devour you.'

⁴Interestingly, *so* is not only used as a correlative, it can also function as a conditional conjunction, similar to French *si* 'if' (< *sic* 'thus, so') (e.g. *So Gott will, wird er wieder gesund* 'If God wants (it), he will get well'). According to Traugott (1985), conditional *so* and *si* were originally used as correlatives that were later extended to subordinate clauses and reanalysed as conditional conjunctions (see also Harris 1986).

Like adverbial clauses, relative clauses may occur with a correlative in the main clause. Linguistic typologists distinguish between several types of relative constructions and one of them is the correlative relative clause (Lipták 2009). Correlative relatives were very frequent in the ancient Indo-European languages (e.g. Hittite, Sanskrit) and are still the dominant relative construction in the Indic branch of modern Indo-European languages (Srivastav 1991). In Hindi, for example, the most frequent type of relative clause is a correlative construction in which the relative clause typically precedes the main clause as in (41).

- (41) Hindi (Indo-European, Indic; Lipták 2009: 1)

[jo laRkii khaRii hai] vo lambii hai.

REL girl standing is that tall is

'The girl who is standing is tall.'

The relative clauses of correlative constructions are non-embedded clauses that typically include the head they modify. In (41), *laRkii* 'girl' is the nominal head of the relative clause, which is marked by the morpheme *jo* and resumed in the second clause by the correlative *vo* 'that'. *Vo* is a case-inflected demonstrative pronoun that is obligatory in this context and serves to indicate the syntactic function of the head within the main clause.

Similar types of correlative relative clauses occur in other languages of our sample. Like the correlative constructions of Hindi (and other Indic languages), the correlative constructions of these languages consist of non-embedded relative clauses in which the nominal head is "represented" by a demonstrative correlative in the main clause. Two examples from Wappo and Georgian are given in (42) and (43).

- (42) Wappo (Wappo-Yukian; Thompson et al. 2006: 115)

ah [i-ø k'ew-ø naw-ta] ce hak'-še?.

1SG.NOM 1SG-ACC man-ACC see-PST.DEP DEM like-DUR

'I like the man I saw.'

- (43) Georgian (Kartvelian; Hewitt 1995: 607)

[gušin rom beč'ed-i (\emptyset - \emptyset)-m-a-čuk-e], is

yesterday SUB ring-NOM you-it-me-LOC-present-AOR.IND that.NOM

(beč'ed-i) sad ar-i-s?

ring-NOM where be-PRS-it

'Where is that ring which you presented to me yesterday?'

Note that the correlative relative clauses in Wappo do not include a marker of the head noun (parallel to Hindi *yo*) and that the correlative constructions in Georgian may include a copy of the head in the second clause (i.e. *beč’ed-i* ‘ring-nom’). In general, correlative relative constructions are very flexible. There is a tendency to prepose the relative clause, but in all of the languages with correlative relatives in our sample, the relative clause may also be postponed to the main clause, and the head noun may occur either within the relative clause (which is most frequent) or in the main clause or in both clauses.

Finally, there are also some languages in our sample in which complement clauses occur with a correlative pronoun. Depending on the order of main and complement clause, the correlative is either forward referring, as in (44) from Hungarian, or it is backwards referring, as in (45) from Thai Kamti. Note that Inglis (2014: 119) refers to the demonstrative in (45) as a “complement marker”, but given that *nai¹* ‘this’ serves as object of the second clause, we consider *nai¹* a backwards referring correlative rather than a complementiser.

- (44) Hungarian (Uralic, Ugric; Kenesei et al. 1998: 28)

Anna tudta (azt), hogy Péter beteg.
Anna knew.DEF that/it.ACC that Peter sick
'Anna knew that Peter was sick.'

- (45) Thai Kamti (Tai-Kadai; Kam-Tai; Inglis 2014: 119)

[tang⁴ man⁴ uu⁵] nai¹ kau³ piuu⁵ uu⁵.
with 3SG live COMP 1SG be.happy IMPF
'(I) am happy that (I) live with her.' (Lit. 'I live with her, this I am happy (about).')

9 Topic markers

The final type of clause linker to be considered in this chapter serves to mark topics. In their basic use, topic markers accompany nominal constituents, but in some languages topic markers also occur with subordinate clauses. For instance, in many languages conditional clauses include a topic marker (Haiman 1985). Haiman (1978) argued that the frequent use of topic markers in conditional clauses is motivated by the communicative function of conditionals to lay the foundation for the interpretation of subsequent clauses (see also Diessel 2005b). But topic markers do not only occur in conditionals; they also appear in other

types of preposed adverbial clauses (Thompson et al. 2007) and certain types of relative clauses (de Vries 1995).

Topic markers are often historically related to copulas and adpositions (e.g. *as for*), but also develop from demonstratives. There are, for instance, several Papuan languages in our sample in which noun phrases, preposed adverbial clauses and internally headed relatives occur with the same demonstrative as topic marker, as in (46) from Usan.

- (46) Usan (Trans-New Guinea, Madang; Reesink 1984: 182, 200, 187)
- a. [munai âib eng] yonou bain mindat-erei.
house big this.GIVEN my older.brother build-3SG.PST
'The big house, my older brother built.'
 - b. [wau eâb igor-iner eng] unor mâni utibâ.
child cry.ss be-3SG.FUT this.GIVEN mother yam she.will.give.him
'If the child is crying, his mother will give him yam.'
 - c. [qemi eng munon bau-or eng] ye me ge-au.
bow this.GIVEN man take-3SG.PST this.GIVEN I not see-NOM
'The bow that the man took I did not see.'

In all three examples, the initial constituent is marked by *eng* 'this.GIVEN', which is composed of the proximal demonstrative *e* 'here, this' and a marker for given information. If *eng* is used as a topic marker, it follows the associated noun phrase or subordinate clause, but *eng* can also function as an independent pronoun meaning 'this/that one' (Himmelmann 1997: 209).

Similar types of topic markers occur in several other Papuan languages (Wambon, Korafe, Menya, Urim). For instance, Wambon (47) and Korafe (48) use demonstratives at the end of preposed subordinate clauses that one might analyse as topics.

- (47) Wambon (Trans-New Guinea, Awju-Dumut; de Vries 1995: 518)
- a. [Wano-ne-e moke-knde-n-eve] kaimo-nde
child-TRS-CONN be.afraid-3PL.PRS-TR-that teacher-CONN
koyomke-khe.
be.angry-3SG.PRS
'The children are afraid because the teacher is angry.'
 - b. [Alive ndu-ne-e takhima-lepo-n-eve] kaimo-nde.
yesterday sago-TR-CONN buy-1SG.PST-TR-DEM good-is
'The sago which I bought yesterday, is good.'

- (48) Korafe (Trans-New Guinea, Binanderean; Farr 1999: 77, 78)

- a. [Nande mandi evetu-fifitu-sira a=mo], jo taima=da
1SG.GEN boy woman-put-PST.3SG that-TOP NEG bush=LOC
sumb-ae=ri.
run-not.do=COP.Q
'When my son got married, he didn't run away with her into the
bush.'
- b. [Nande mandi evetu-fifitu-sira a=mo], oroko Moresby
1SG.GEN boy woman-put-PST.3SG that-TOP today Moresby
ir-ira.
remain-PRS.3SG
'My son that's married is living in Moresby now.'

On the face of it, the subordinate clauses in these examples look similar to some of the nominalised clauses that we have seen in §3. In particular, Jamul Tiipay and Assiniboine have relative and complement clauses that end with a demonstrative (cf. 8c and 9c), but in contrast to the clause-final demonstratives of Usan, Wambon and Korafe, the clause-final demonstratives of Jamul Tiipay and Assiniboine do not occur in conditional clauses and do not seem to serve as topic markers (according to our sources).

Since there are no diachronic corpora to study the development of demonstrative topicalisers, we cannot be certain how these markers have evolved. Yet, Reesink (1984) and de Vries (1995) proposed a scenario which, we believe, provides a plausible account for their development. Both scholars observe that topicalised subordinate clauses in Papuan languages are (often) resumed by a correlate pronoun at the beginning of the main clause, as in (49) from Wambon.

- (49) Wambon (Trans-New Guinea, Awju-Dumut; de Vries 1995: 517, 518)

- a. [Ko mba-khe-n-o kav=eve] eve na-mbap-nde.
there stay-3SG.PRS-TR-CONN man=that that my-father-is
'The man who is staying there, that is my father.'
- b. [Kikhuve ndethekhel=eve] eve Manggelum konoksiva.
Digul rise.3SG.COND=that that Manggelum go.NEG.1PL.INTENT
'If the Digul river rises, then we do not want to go to Manggelum.'

According to de Vries, the demonstrative topicalisers of Wambon (and other Papuan languages) are derived from demonstrative correlatives that have become associated with the preceding subordinate clause. There is good evidence

for this hypothesis, especially in Wambon. Since *eve* ‘that’ is a demonstrative pronoun that cannot be interpreted as a determiner if it follows an NP or clause (demonstrative determiners precede NPs in Wambon), it seems reasonable to assume that *eve* evolved from a correlate pronoun rather than from a nominalising article, or determiner, as some of the demonstrative clause linkers described in §3 (Reesink 1984: 187–188).

10 Summary and conclusion

To conclude, demonstratives are of fundamental importance to clause combining. They are commonly used as anaphors and discourse deictics and provide a very frequent source for the development of various types of grammatical clause linkers. Some of these developments are frequently mentioned in textbooks and handbook chapters on grammaticalisation, but others have only been described in reference grammars and other special sources. The current study provides the first large-scale investigation of demonstrative clause linkers from a cross-linguistic perspective. Drawing on data from a sample of 100 languages, the chapter has analysed eight basic types of clause linkers that are frequently derived from a demonstrative:

1. Relative pronouns
2. Linking and nominalising articles
3. Quotative markers
4. Complementisers
5. Conjunctive adverbs
6. Adverbial subordinate conjunctions
7. Correlatives
8. Topic markers

There is abundant evidence in our database that all of these markers are often etymologically related to demonstratives. Yet, while the deictic origins of many clause linkers are morphologically transparent, it is not always clear how they evolved. In accordance with the current literature on grammaticalisation, we have argued that the development of demonstratives into grammatical clause

linkers is crucially influenced by the constructions in which demonstratives occur. If we want to understand how and why demonstratives develop into grammatical clause linkers, we must not only consider the discourse-pragmatic uses of demonstratives but also their syntactic functions.

One aspect that is not always recognised in the literature on grammaticalisation is that not all demonstrative clause linkers are immediately derived from demonstrative anaphors and discourse deictics. As we have seen, the various types of demonstrative clause linkers are historically related to each other, and these relationships are crucial for understanding the occurrence of demonstratives in certain clause-linkage constructions. In particular, the analysis of demonstratives in subordinate clauses needs to take into account that the various types of subordinate markers are historically related (Schmidtke-Bode 2014; Diessel 2019b). For instance, contrary to what is commonly assumed in the literature (e.g. Hopper & Traugott 2003: 184–185), the demonstratives of nominal and adverbial clauses are often based on demonstrative relative markers and articles rather than on demonstrative clause linkers of paratactic sentences. While there are languages in which complementisers and adverbial subordinators are immediately derived from the discourse uses of demonstratives (see McConvell 2006 for some examples from Australian languages), this does not seem to be a major path of evolution.

Finally, on a more general note, this chapter presents new evidence for Brugmann's (1904) and Bühler's (1934) claim that many grammatical function morphemes have a deictic origin. Current research on grammaticalisation has been mainly concerned with the development of grammatical markers from content words and has paid little attention to demonstratives. In fact, some researchers have argued that all grammaticalisation processes evolve from nouns and verbs (Heine & Kuteva 2007: 111). However, in addition to nouns and verbs, demonstratives provide an extremely frequent source for the development of a wide range of grammatical markers, including the many different types of clause linkers investigated in the current study. More research is needed to better understand the mechanisms behind some of the developments considered in this chapter. Yet, there is no doubt that demonstratives are of fundamental significance to the diachronic evolution of grammar, including the evolution of grammatical clause linkers and clause linkage constructions.

Abbreviations

The chapter abides by the Leipzig Glossing Rules. Additional or deviant abbreviations include:

ACT	actor voice	MID	middle voice
AOR	aorist	NARR	narrative
ASP	aspect	NHYP	nonhypothetical
ASP.NEG	aspect negation	OUT	V outward
AUG	augment	PAT	patient
CONN	connective	POT	potential, hypothetical
CT	contrastive	PRE	prefix
DP	dependent marker	PRES	presentative
EMPH	emphatic	PRN	pronoun
EP	epenthetic	RPST	remote past tense
EVD	indirect evidence	RC	relative construction
GIVEN	given information	SJREL	subject relative
HAB	habitual	SS	same subject
IMPF	imperfect	ST	first part of a discontinuous root
INAN	inanimate	START	inceptive
INTENT	intentional	SUB	subordinator
LK	linker/linking article	TRS	transitional sound
MED	medial demonstrative		

Appendix: Language sample

AFRICA: Fongbe, Hausa, Jamsay, Kana, Khwe, Konso, Koyra Chiini, Krongo, Lango, Mayogo, Mbay, Nkore Kiga, Noon, Supyire, Tamashék

NORTH AND CENTRAL AMERICA: Assiniboine, Choctaw, (Barbareño) Chumash, Jamul Tiipay, Kiowa, (Chalcatongo) Mixtec, Musqueam, Ojibwe, Purépecha, Rama, Slave, Tepehua, Tümpisa Shoshone, Tzutujil, Wappo, West Greenlandic

SOUTH AMERICA: Aguaruna, Awa Pit, Barasano, Cavineña, Epena Pedee, Hixkaryana, Hup, Jarawara, Kwazá, Mapudungun, Matsés, Mekens, Mosetén, (Hual-laga) Quechua, Tariana, Trumai, Urarina, Warao, Wari', Yagua, Yurakaré

EURASIA: Abkhaz, Arabic, Basque, Evenki, French, Georgian, German, Hindi, Hungarian, Japanese, Korean, Lezgian, Malayalam, Santali, Serbo-Croatian, Turkish, (Kolyma) Yukaghir

SOUTHEAST ASIA AND OCEANIA: Burmese, Hmong Njua, Begak Ida'an, Toba Batak, Lao, Mandarin Chinese, Meithei, Semelai, Tagalog, Thai Kamti, Tetun, Toqabaqita, Tukang Besi, Vietnamese

AUSTRALIA AND NEW GUINEA: Alamblak, Ambulas, Amele, Bariai, Kayardild, Komnzo, Korafe, Manambu, Mangarayi, Martuthunira, Menya, Motuna, Ungarinjin, Usan, Wambaya, Wambon, Yimas

References

- Aikhenvald, Alexandra Y. 2008. *The Manambu language of East Sepik, Papua New Guinea*. Oxford: Oxford University Press.
- Aristar, Anthony Rodrigues. 1991. On diachronic sources and synchronic pattern: An investigation into the origin of linguistic universals. *Language* 67. 1–33.
- Axel, Katrin. 2009. Die Entstehung des *dass*-Satzes – ein neues Szenario. In Veronika Ehrich, Christian Fortmann & Marga Reise (eds.), *Koordination und Subordination im Deutschen*, 21–41. Hamburg: Buske.
- Behaghel, Otto. 1928. *Deutsche Syntax: Eine geschichtliche Darstellung*. Vol. 3: *Satzgebilde*. Heidelberg: Winter.
- Bruce, Les. 1984. *The Alamblak language of Papua New Guinea (East Sepik)*. Canberra: Pacific Linguistics.
- Brugmann, Karl. 1904. *Die Demonstrativpronomina der indogermanischen Sprachen*. Leipzig: Teubner.
- Bühler, Karl. 1934. *Sprachtheorie: Die Darstellungsfunction der Sprache*. Jena: Fischer.
- Chelliah, Shobhana L. 1997. *A grammar of Meithei*. Berlin: Mouton de Gruyter.
- Comrie, Bernard. 2006. Syntactic typology: Just how exotic ARE European-type relative clauses? In Ricardo Mairal (ed.), *Linguistic universals*, 130–154. Cambridge: Cambridge University Press.
- Comrie, Bernard. 2013. Alignment of case marking of full noun phrases. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/98>.
- Coventry, Kenny R., Berenice Valdés, Alejandro Castillo & Pedro Guijarro-Fuentes. 2008. Language within your reach: Near–far perceptual space and spatial demonstratives. *Cognition* 108(3). 889–895.
- Cumberland, Linda A. 2005. *A grammar of Assiniboine, a Siouan language of the Northern Plains*. Bloomington, IN: Indiana University. (PhD dissertation).
- Cyr, Danielle E. 1993. Demonstratives and definite articles in Plains Cree. In William Cowan (ed.), *Papers from the Twenty-Fourth Algonquian Conference*, 64–80. Ottawa: Carleton University.
- Dayley, John. 1989. *Tümpisa (Panamint) Shoshone grammar*. Berkeley, CA: University of California at Berkeley.

- de Vries, Lourens. 1995. Demonstratives, referent identification and topicality in Wambon and some other Papuan languages. *Journal of Pragmatics* 24(5). 513–533.
- Deutscher, Guy. 2000. *Syntactic change in Akkadian: The evolution of sentential complementation*. Oxford: Oxford University Press.
- Deutscher, Guy. 2001. The rise and fall of a rogue relative construction. *Studies in Language* 25. 405–422.
- Deutscher, Guy. 2009. Nominalization and the origin of subordination. In Talmy Givón & Masayoshi Shibatani (eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution*, 199–214. Amsterdam: Benjamins.
- Diessel, Holger. 1999a. *Demonstratives: Form, function, and grammaticalization*. Amsterdam: Benjamins.
- Diessel, Holger. 1999b. The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology* 3(1). 1–49.
- Diessel, Holger. 2005a. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43. 449–470.
- Diessel, Holger. 2005b. Pronominal and adnominal demonstratives. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *World atlas of language structures*, 174–179. Oxford: Oxford University Press.
- Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
- Diessel, Holger. 2012. Bühler's two-field theory of pointing and naming and the deictic origins of grammatical morphemes. In Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and language change: New reflections*, 37–50. Amsterdam: Benjamins.
- Diessel, Holger. 2013. Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. *Language and Cognition* 5(2–3). 239–249.
- Diessel, Holger. 2014. Demonstratives, frames of reference, and semantic universals of space. *Language and Linguistics Compass* 8(3). 116–132.
- Diessel, Holger. 2019a. Preposed adverbial clauses: Functional adaptation and diachronic inheritance. In Karsten Schmidke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja A. Seržant (eds.), *Explanation in linguistic typology: Diachronic sources, functional motivations and the nature of the evidence*, 97–122. Berlin: Language Science Press.
- Diessel, Holger. 2019b. *The grammar network: How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.
- Döhler, Christian. 2018. *A grammar of Komizo*. Berlin: Language Science Press.

- Dryer, Matthew S. 1989. Article-noun order. In Randolph Graczek, Brad Music & Caroline Wiltshire (eds.), *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society*, 83–97. Chicago, IL: Chicago Linguistic Society.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138.
- Ebert, Robert Peter. 1978. *Historische Syntax des Deutschen*. Stuttgart: Metzler.
- Farr, Cynthia J. M. 1999. *The interface between syntax and discourse in Korafe*. Canberra: Pacific Linguistics.
- Foley, William A. 1980. Toward a universal typology of the noun phrase. *Studies in Language* 4(2). 171–199.
- Foley, William A. 1991. *The Yimas language of New Guinea*. Stanford, CA: Stanford University Press.
- Gallagher, Steve & Peirce Baehr. 2005. *Bariai grammar sketch*. Ukarumpa: Summer Institute of Linguistics.
- Givón, Talmy. 1991. The evolution of dependent clause morphosyntax in Biblical Hebrew. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. 2: Types of grammatical markers, 257–310. Amsterdam: Benjamins.
- Givón, Talmy. 2009. *The genesis of syntactic complexity*. Amsterdam: Benjamins.
- Golato, Andrea. 2000. An innovative German quotative for reporting on embodied actions: *Und ich so/und er so* ‘and I’m like/and he’s like’. *Journal of Pragmatics* 32(1). 29–54.
- Guérin, Valérie (ed.). 2019. *Bridging constructions*. Berlin: Language Science Press.
- Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.
- Haiman, John. 1978. Conditionals are topics. *Language* 54(3). 564–589.
- Haiman, John. 1985. *Natural syntax*. Amsterdam: Benjamins.
- Haiman, John & Pamela Munro (eds.). 1983. *Switch reference and universal grammar. Proceedings of a symposium on switch reference and universal grammar, Winnipeg, May 1981*. Amsterdam: Benjamins.
- Halliday, M. A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Routledge.
- Harms, Phillip Lee. 1994. *E pena Pedee syntax*. Dallas, TX: The Summer Institute of Linguistics & The University of Texas at Arlington.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Harris, Martin. 1978. *The evolution of French syntax: A comparative approach*. London: Longman.

- Harris, Martin. 1986. The historical development of conditional sentences in Romance. *Romance Philology* 39. 405–436.
- Heath, Jeffrey. 2005. *A grammar of Tamashk*. Berlin: Mouton de Gruyter.
- Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
- Hendery, Rachel. 2013. Constructional etymology: The sources of relative clauses. In Robert Mailhammer (ed.), *Lexical and structural etymology*, 83–119. Berlin: Mouton de Gruyter.
- Hetzron, Robert. 1995. Genitival agreement in Awngi: Variation on an Afroasiatic theme. In Frans Plank (ed.), *Double case: Agreement by Suffixaufnahme*, 325–335. Oxford: Oxford University Press.
- Hewitt, Brian George. 1995. *Georgian: A structural reference grammar*. Berlin: Mouton de Gruyter.
- Himmelmann, Nikolaus. 1997. *Deiktikon, Artikel, Nominalphrase. Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.
- Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.
- Inglis, Douglas Maxton. 2014. *This here thing: Specifying morphemes an³, nai¹, and mai² in Tai Khamti reference point constructions*. Edmonton: University of Alberta. (PhD dissertation).
- Kenessei, István, Robert M. Vago & Anna Fenyvesi. 1998. *Hungarian*. London: Routledge.
- Klamer, Marian. 2000. How report verbs become quote markers and complementisers. *Lingua* 110(2). 69–98.
- König, Ekkehard. 2012. Le rôle des déictiques de manière dans le cadre d'une typologie de la déixis. *Bulletin de la Société de Linguistique de Paris* 107(1). 11–42.
- Laury, Ritva. 1997. *Demonstratives in interaction: The emergence of a definite article in Finnish*. Amsterdam: Benjamins.
- Lehmann, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Narr.
- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam: Benjamins.
- Lipták, Anikó. 2009. The landscape of correlatives. An empirical and analytical survey. In Anikó Lipták (ed.), *Correlatives cross-linguistically*, 1–46. Amsterdam: Benjamins.
- Lockwood, William B. 1968. *Historical German syntax*. Oxford: Clarendon.

- Lord, Carol. 1993. *Historical change in serial verb constructions*. Amsterdam: Benjamins.
- Lühr, Rosemarie. 2008. Competitive Indo-European syntax. In Gisella Ferraresi & Maria Goldbach (eds.), *Principles of syntactic reconstruction*, 121–159. Amsterdam: Benjamins.
- McConvell, Patrick. 2006. Grammaticalization of demonstratives as subordinate complementizers in Ngumpin-Yapa. *Australian Journal of Linguistics* 26(1). 107–137.
- Miller, Amy. 2001. *A grammar of Jamul Tiipay*. Berlin: Mouton de Gruyter.
- Neukom, Lukas. 2001. *Santali*. Munich: Lincom.
- Obata, Kazuko. 2003. *A grammar of Bilua. A Papuan language of the Solomon Islands*. Canberra: Pacific Linguistics.
- Paul, Hermann. 1916–1920. *Deutsche Grammatik*. Halle (Saale): Niemeyer.
- Payne, Doris L. & Thomas E. Payne. 1990. Yagua. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 2, 249–471. Berlin: Mouton de Gruyter.
- Pittner, Karin. 1995. The case of German relatives. *The Linguistic Review* 12(3). 197–231.
- Post, Mark William. 2007. *A grammar of Galo*. Bundoora, VIC: La Trobe University. (Doctoral dissertation).
- Radford, Andrew. 1997. *Syntactic theory and the structure of English. A minimalist approach*. Cambridge: Cambridge University Press.
- Reesink, Gerard Pieter. 1984. *Structures and their function in Usan. A Papuan language of Papua New Guinea*. Amsterdam: University of Amsterdam. (PhD dissertation).
- Roberts, John R. 1987. *Amele*. London: Croom Helm.
- Ross, Malcolm D. 1988. *Proto Oceanic and the Austronesian languages of Western Melanesia*. Canberra: Pacific Linguistics.
- Schmidtke-Bode, Karsten. 2014. *Complement clauses and complementation systems: A cross-linguistic study of grammatical organization*. Jena: Friedrich-Schiller-Universität Jena. (PhD dissertation).
- Schmidtke-Bode, Karsten & Holger Diessel. 2017. Cross-linguistic patterns in the structure, function, and position of (object) complement clauses. *Linguistics* 55(1). 1–38.
- Schuh, Russell G. 1983. The evolution of determiners in Chadic. In Ekkehard Wolff & Hilke Meyer-Bahlburg (eds.), *Studies in Chadic and Afroasiatic linguistics*, 157–210. Hamburg: Buske.

- Schuh, Russell G. 1990. Re-employment of grammatical morphemes in Chadic: Implications for language history. In Philip Baldi (ed.), *Linguistic change and reconstruction methodology*, 599–618. Berlin: Mouton de Gruyter.
- Sohn, Ho-Min. 1994. *Korean*. London: Routledge.
- Sohn, Ho-Min. 2009. The semantics of clause linkage in Korean. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The semantics of clause linkage. A cross-linguistic typology*, 285–317. Oxford: Oxford University Press.
- Soukka, Maria. 2000. *A descriptive grammar of Noon*. Munich: Lincom.
- Srivastav, Veneeta. 1991. The syntax and semantics of correlatives. *Natural Language and Linguistic Theory* 9(4). 637–686.
- Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang. 2007. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, 2nd ed, vol. 2: *Complex constructions*, 237–300. Cambridge: Cambridge University Press.
- Thompson, Sandra A., Joseph Sung-Yul Park & Charles N. Li. 2006. *A reference grammar of Wappo*. Berkeley, CA: University of California Press.
- Traugott, Elizabeth Closs. 1985. Conditional markers. In John Haiman (ed.), *Iconicity in syntax*, 289–307. Amsterdam: Benjamins.
- Traugott, Elizabeth Closs. 2003. Constructions in grammaticalization. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 624–647. Malden, MA: Blackwell.
- Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- van der Auwera, Johan. 1985. Relative *that* – a centennial dispute. *Journal of Linguistics* 21(1). 149–179.
- van Gijn, Rik. 2006. *A grammar of Yurakaré*. Nijmegen: Radboud University. (Doctoral dissertation).
- Wash, Suzanne. 1999. On the structure and function of relative clauses in Barbareño Chumash. Santa Barbara, CA: University of California at Santa Barbara. Unpublished ms.
- Wash, Suzanne. 2001. *Adverbial clauses in Barbareño Chumash narrative discourse*. Santa Barbara, CA: University of California at Santa Barbara. (PhD dissertation).
- Webber, Bonnie Lynn. 1991. Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes* 6(2). 107–135.
- Wilson, Patricia R. 1980. *Ambulas grammar*. Ukarumpa: Summer Institute of Linguistics.

Name index

- Abraham, Werner, 139, 186, 189, 193, 201
Acton, Eric K., 12
Ahrenholz, Bernt, 187, 188, 190, 194, 207, 208
Aijmer, Karin, 139
Aikhenvald, Alexandra Y., 94, 174, 324
Anderson, Stephen R., 13, 104, 244
Arbib, Michael A., 22
Ariel, Mira, 132, 139, 187, 244
Aristar, Anthony Rodrigues, 312–314
Arkhangelskiy, Timofey, 288
Auer, Peter, 39
Axel, Katrin, 319, 320
- Baehr, Peirce, 315
Barbour, Julie, 75
Barske, Tobias, 30
Bates, Douglas, 198
Behaghel, Otto, 319
Benus, Stefan, 137
Betz, Emma, 30
Beznosikova, Ljutsija M., 285
Bisle-Müller, Hansjörg, 188
Bliss, Heather, 125–127, 129, 133, 134
Bolinger, Dwight, 35
Bosch, Peter, 187, 191, 192, 196, 204, 206
Boulonnais, Dominique, 35
Brennan, Susan E., 105
Bril, Isabelle, 10
- Brotchie, Amanda, 86
Brown, Cheryl, 188
Brown, Penelope, 219, 220
Bruce, Les, 323
Brugmann, Karl, 23, 306–308, 334
Buchstaller, Isabelle, 273
Bühler, Karl, 25, 306–308, 334
Burenhult, Niclas, 1, 105
Bürkner, Paul-Christian, 203
Burridge, Kate, 2
- Campbell, Lyle, 310, 319
Carol, Javier, 180
Carpio, María Belén, 150–153, 178
Chafe, Wallace, 116
Chelliah, Shobhana L., 316
Chiriacescu, Sofiana Iulia, 189
Clark, Herbert H., 105, 274, 293
Colasanti, Valentina, 132
Comrie, Bernard, 187, 308, 312
Coventry, Kenny R., 244, 306
Croft, William, 220
Crowley, Terry, 72, 92, 94
Cruse, D. Alan, 220
Crystal, David, 279
Cumberland, Linda A., 313
Cúneo, Paola, 150, 151, 153, 174
Cyr, Danielle E., 306
- Dahlstrom, Amy, 133
Dawuda, Carmen, 1, 5
Dayley, John, 309

- de Vries, Lourens, 86, 331, 332
Degand, Liesbeth, 2–4, 6, 7, 31, 33, 39
Deichsel, Annika, 189
Delsing, Lars-Olof, 219
Deutscher, Guy, 313, 314
Diessel, Holger, 1, 4, 5, 10, 13, 22, 23,
 43, 80, 82, 85, 104–106, 131,
 151, 152, 155, 156, 172, 174,
 216, 220–222, 237, 244, 245,
 262, 305–307, 311, 312, 318,
 325, 330, 334
Diewald, Gabriele, 31
Dixon, R. M. W., 4, 22, 104, 244
Döhler, Christian, 317
Dömötör, Adrienne, 293, 294
Dostie, Gaétane, 2, 3
Dowty, David, 187, 191
Dryer, Matthew S., 307, 311
Duvallon, Outi, 251

Early, Robert J., 86, 87, 89
Ebert, Robert Peter, 319
Echols, John M., 117
Eckardt, Regine, 52
Edygarova, Svetlana, 277
Ellert, Miriam, 187
Enfield, Nick, 1, 244
Englebretson, Robert, 2
Erelt, Mati, 275
Etelämäki, Marja, 250, 251, 260, 261
Evans, Nicholas, 4, 11, 222
Evers-Vermeul, Jacqueline, 31, 39

Faarlund, Jan Terje, 218
Fagard, Benjamin, 33
Fairbanks, Brendan, 133, 139
Farr, Cynthia J., 11
Farr, Cynthia J. M., 315, 332
Fedjunëva, Galina V., 286

Fillmore, Charles J., 104
Flatty, Don A. L., 118
Florey, Margaret, 2
Foley, William A., 311, 312, 326, 327
Frantz, Donald G., 124, 126, 139–142
Fraser, Bruce, 3

Gallagher, Steve, 315
Garbacz, Piotr, 217
Garrod, Simon, 192
Geluykens, Roland, 262
Genee, Inge, 125, 133, 139
Georgakopoulos, Thanasis, 63
Gerner, Matthias, 22
Gernsbacher, Morton Ann, 188, 189,
 191, 192, 199
Gerrig, Richard J., 274, 293
Gildea, Spike, 174
Givón, Talmy, 188, 192, 193, 310, 320
Goddard, Ives, 133
Golato, Andrea, 30, 316
González, Raúl Eduardo, 150, 151, 153
Grenoble, Lenore A., 2, 249, 256
Grice, H. Paul, 48
Gudde, Harmen B., 244
Guérin, Valérie, 44, 86, 323
Güldemann, Tom, 273–276, 284, 298,
 314, 317
Gundel, Jeanette K., 187, 248, 265
Gust, Helmar, 24
Gutiérrez, Analía, 174, 175, 180

Hagen, Kristin, 223, 224
Haider, Hubert, 196
Haiman, John, 324, 330
Hakulinen, Auli, 251, 252, 275, 280
Halliday, M. A. K., 105, 245, 306
Hanks, William F., 1, 105, 116, 244
Hansen, Maj-Britt Mosegaard, 2

Name index

- Hanssen, Eskil, 224
Harms, Phillip Lee, 317
Harris, Alice C., 310, 319
Harris, Martin, 306, 328
Hasan, Ruqaiya, 105, 245, 306
Haspelmath, Martin, 70
Heath, Jeffrey, 310, 318, 326, 327
Heine, Bernd, 2–4, 306, 310, 326, 334
Helasvuo, Marja-Liisa, 277
Hellwig, Birgit, 180
Hendery, Rachel, 310
Hetzron, Robert, 312
Hewitt, Brian George, 328, 329
Himmelmann, Nikolaus, 1, 5, 10, 13,
 85, 87, 97, 105, 114, 116, 154,
 172, 178–180, 185, 221, 222,
 244, 306, 307, 311, 312, 331
Hinterwimmer, Stefan, 187, 191
Hopper, Paul J., 326, 334
Hyslop, Catriona, 86

Inglis, Douglas Maxton, 316, 330
Itkonen, Terho, 250, 252

Janda, Laura, 63
Jauncey, Dorothy, 86
Jefferson, Gail, 248, 267
Johannessen, Janne Bondi, 187, 215–
 218, 222–225, 236–238
Jónsson, Jóhannes Gísli, 135
Jucker, Andreas H., 33
Junker, Marie-Odile, 133
Jurafsky, Dan, 62, 63

Kaiser, Elsi, 187
Karssenberg, Lena, 44, 299
Keenan, Edward L., 13, 104, 244
Keevallik, Leelo, 32, 275, 281, 298
Kel'makov, Valentin K., 277

Kemmerer, David, 104
Kenesei, István, 308, 328, 330
Kibrik, Andrej A., 22, 244
Kiefer, Ferenc, 275, 295, 296
Klamer, Marian, 314
Klein, Harriet, 152, 153, 174, 175
König, Ekkehard, 2, 7, 22–24, 44, 56,
 57, 274, 280, 284, 298, 299,
 314, 315, 321, 326
Körtvély, Erika, 295, 296
Krasnoukhova, Olga, 155, 180
Kretzschmar, Franziska, 192
Küntay, Aylin C., 1
Kuteva, Tania, 306, 310, 326, 334

Labov, William, 61
Lahousse, Karen, 44, 299
Lakoff, George, 63
Lakoff, Robin, 11, 221
Lam, Zoe Wai-Man, 133
Lange, Annalena, 188, 205
Larjavaara, Matti, 250, 254
Laury, Ritva, 245, 249–252, 260, 261,
 264, 265, 306
Lehmann, Christian, 308, 311
Leinonen, Marja, 291
Levinson, Stephen C., 13, 105, 180,
 219, 220, 245, 246
Lewandowska-Tomaszczyk, Barbara, 63
Lichtenberk, Frantisek, 5, 86
Lie, Svein, 215, 220–223, 229, 237
Lipták, Anikó, 327, 329
Lockwood, William B., 310, 327
Lord, Carol, 314, 320
Lühr, Rosemarie, 319, 320
Luís, Ana, 77, 82
Lyons, John, 29, 104, 105

- Maes, Alfons, 245
March, Evrim G., 244
Margetts, Anna, 2, 5, 10
Maschler, Yael, 39
Mayer, Mercer, 114
Mazzon, Gabriella, 2
McConvell, Patrick, 334
Meira, Sérgio, 13, 244
Messineo, Cristina, 150–153, 157, 174
Metslang, Helle, 275
Miller, Amy, 313
Müller, Christoph, 197
Munro, Pamela, 324

Nercesian, Verónica, 180
Neukom, Lukas, 322
Nikitina, Tatiana, 54, 59, 62
Nishina, Yoko, 22, 23, 44
Nøklestad, Anders, 224
Nordlinger, Rachel, 174

Obata, Kazuko, 323
Östman, Jan-Ola, 252
Oxford, Will, 134
Özden, Filiz, 188, 205
Özyürek, Aslı, 1, 105

Pajusalu, Renate, 250, 258
Parker, Gary J., 73, 76
Patil, Umesh, 188, 205, 209
Pattison, Philippa, 244
Paul, Hermann, 310
Payne, Doris L., 153, 309
Payne, Thomas E., 309
Pischlöger, Christian, 277
Pittner, Karin, 310
Polis, Stéphane, 63
Post, Mark William, 323
Potts, Christopher, 11, 12

Primus, Beatrice, 187, 191
Prince, Ellen F., 207
Pryor, John, 11
Pullum, Geoffrey K., 77, 82

Radford, Andrew, 317
Reesink, Gerard Pieter, 298, 299, 316,
 331–333
Reile, Maria, 245, 246, 248, 250, 257,
 258
Rett, Jessica, 137
Ridge, Eleanor, 71, 74, 94
Riley, Matthew, 2
Rishøi, Ingvild H., 224, 226–232
Roberts, John R., 318
Rooij, Christ, 245
Ross, Malcolm D., 312
Russell, Lena, 125, 139
Russell, Norma Jean, 139–142

Sadler, Louisa, 174
Sahkai, Heete, 249
San Roque, Lila, 222
Sanford, Tony, 192
Schapper, Antoinette, 118, 222
Schiffrin, Deborah, 2, 3, 132, 139
Schmidtke-Bode, Karsten, 318–321,
 334
Schneider, Cindy, 86
Schuh, Russell G., 312–314
Schumacher, Petra B., 187, 188, 190–
 192, 196, 204, 206, 209
Schupbach, Shannon Scott, 126, 127
Schwarz, Florian, 11
Seppänen, Eeva-Leena, 250
Sheljakin, M. A., 249
Shor, Leon, 44
Shroyer, Suzanne, 188, 189, 191, 192,
 199

Name index

- Sichel, Ivy, 143
Siitonen, Kirsti, 252
Silverstein, Michael, 247
Šimík, Radek, 11
Simon-Vandenbergen, Anne-Marie,
 2–4, 6, 7
Slobin, Dan I., 244
Sneddon, James Neil, 117
Sohn, Ho-Min, 324
Soukka, Maria, 320
Spencer, Andrew, 77, 82
Srivastav, Veneeta, 329
Stevenson, Rosemary J., 209
Strahan, Tania, 220
Strube, Michael, 197

Talmy, Leonard, 23, 252
Teleman, Ulf, 216, 219, 220, 223
Teptiuk, Denys, 278, 279
Terrill, Angela, 13, 244
Thieberger, Nicholas, 86
Thoma, Sonja, 133, 142, 143
Thompson, Chad, 143
Thompson, Sandra A., 325, 329, 331
Timberlake, Alan, 249
Tomasello, Michael, 23
Traugott, Elizabeth Closs, 34, 307,
 326, 328, 334
Treis, Yvonne, 13, 48
Trousdale, Graeme, 307
Trueswell, John C., 187

Umbach, Carla, 24, 44, 187, 274
Uotila, T. E., 277, 286, 290
Usher, Timothy, 118

van Alphen, Ingrid, 273
van der Auwera, Johan, 308
Van Der Mark, Sheena, 126

van Gijn, Rik, 322
Vezzosi, Letizia, 23
Vidal, Alejandra, 150–153, 174–176
Vindenes, Urd, 229
Visser, Eline, 104
von Heusinger, Klaus, 187–189
von Prince, Kilu, 72

Waletzky, Joshua, 61
Wash, Suzanne, 312
Webber, Bonnie Lynn, 306
Weinert, Regina, 207, 208
Weinrich, Harald, 186, 189, 193, 194,
 207, 208
Whitehead, Carl R., 11
Wiese, Heike, 31
Wilson, Patricia R., 315
Wiltschko, Martina, 132, 143

Zifonun, Gisela, 188, 192, 204
Zimmermann, Malte, 133, 139
Zubova, Yulia N., 288
Zwickly, Arnold M., 35, 77, 82

Language index

- Abkhaz, 335
Afrikaans, 187
Afro-Asiatic, 313
Aguaruna, 335
Akkadian, 313, 314
Alaaba, 13
Alamblak, 323, 324, 336
Algonquian, 123, 133, 133⁸
Ambulas, 314, 315, 336
Amele, 318, 336
Arabic, 335
Armenian, 26
Assiniboine, 313, 318, 332, 335
Austronesian, 69, 117, 311, 312
Awa Pit, 323, 335
- Babylonian, 313
Barasano, 335
Bariai, 315, 336
Basque, 335
Begak Ida'an, 335
Bilinarra, 326
Bilua, 323
Bislama, 71, 75³
Blackfoot, 9–11, 11⁴, 13, 15, 123, 123¹, 124, 124², 125–129, 131–134, 137–140, 142, 143
Botin, 11
Bunaq, 104, 118, 118⁷, 119, 119⁸
Burmese, 323, 335
Cantonese, 132, 133
- Cavineña, 335
Chadic, 312
Chinese, *see* Cantonese, *see* Mandarin
Choctaw, 335
Chorote, 180
Chumash, 312, 318, 323, 335
Cushitic, 44, 312
Czech, 11
- Daakaka, 72
Danish, 217
Dene, 143
Diegueño, 318
Dutch, 35, 187
- English, 8, 9, 11, 22–29, 29⁶, 30, 30⁷, 32–39, 57, 71, 75³, 131, 132, 137, 139, 155, 156, 188, 189, 276, 281¹³, 306, 314, 318, 319, 321, 322, 325–327
Epena Pedee, 317, 323, 335
Estonian, 7, 8, 16, 32⁹, 44¹, 245, 245², 246, 247, 247⁵, 248–250, 253, 254, 256–259, 265, 266, 274, 275, 277, 279–285, 297, 298, 300
Evenki, 335
- Finnic, 248, 274
Finnish, 7, 8, 16, 23, 56⁶, 245, 245², 246, 247, 247⁵, 248–250,

Language index

- 250⁶, 251–254, 254⁹, 259, 260, 262, 264–266, 274, 275, 277, 278, 278⁹, 279–285, 293, 297, 298, 321
Finno-Ugric, 5, 8, 16, 274, 275, 297, 298
Foe, 11³
Fongbe, 335
French, 2, 3², 22, 26, 27, 27⁴, 28, 28⁵, 33, 37, 44¹, 132, 299, 308, 315, 326, 328⁴, 335
Galo, 322
Gamo, 13
Georgian, 308, 328–330, 335
German, 7, 11, 15, 22–27, 27³, 27⁴, 28–30, 32, 33, 35, 37, 39, 52, 132, 133, 143, 185–192, 194, 195, 198⁵, 205, 207, 209, 284, 299, 308–310, 314–320, 322, 323, 325–328, 335
Middle High German, 319, 320
Old High German, 310
Germanic, 25, 35
Goemai, 180
Guaykuruan, 149, 150⁵, 151, 152, 171, 174, 179, 180
Hausa, 335
Hebrew, 44¹, 143
Hindi, 329, 330, 335
Hittite, 329
Hixkaryana, 323, 335
Hmong Njua, 335
Hungarian, 8, 16, 56⁶, 274, 275, 277, 279, 284, 285, 292, 292¹⁷, 293–300, 308, 308¹, 327, 328, 330, 335
Hup, 335
Icelandic, 135, 217
Ilokano, 311
Indic, 329
Indo-European, 143, 329
Indonesian, 104, 110², 117–119
Italian, 22, 26–28, 37, 132
Italo-Romance, 132
Jahai, 105
Jamsay, 335
Jamul Tiipay, 312, 312², 313, 318, 326, 332, 335
Japanese, 11, 22, 23, 24², 28, 44¹, 132, 321, 323, 335
Jarawara, 335
Kalamang, 7, 10, 13, 15, 103, 104, 106, 107, 109, 115, 117, 118, 118⁷, 119
Kambaata, 8, 13, 14, 44, 45, 45², 46–48, 48³, 49–59, 61–63, 65
Kana, 335
Kayardild, 336
Khwe, 335
Kiowa, 335
Komi, 8, 16, 274, 275, 277, 278, 278⁸, 279, 285, 285¹⁴, 286, 286¹⁵, 289–292, 297, 298, 300
Komnzo, 317, 336
Konso, 335
Korafe, 11, 315, 324, 331, 332, 336
Korean, 324, 335
Koyra Chiini, 323, 335
Kongo, 335
Kwazá, 335
Lakhota, 318
Lango, 335
Lao, 335

- Latin, 13, 33
Lewo, 87, 89
Lezgian, 335

Malayalam, 335
Manambu, 323, 324, 336
Mandarin, 335
Mande, 44, 59⁷
Mangarayi, 336
Mapudungun, 335
Martuthunira, 336
Matsés, 335
Mayogo, 335
Mbaham, 118
Mbay, 335
Medumba, 132
Meithei, 316, 335
Mekens, 335
Menya, 323, 324, 331, 336
Meskwaki, 133
Mixtec, 335
Mosetén, 335
Motuna, 336
Movima, 180
Musqueam, 335

Nama, 13⁵
Neverver, 75
Nivaçle, 180
Nkore Kiga, 335
Noon, 320, 335
Norwegian, 10–12, 16, 187, 216, 217, 224, 234⁸

Oceanic, 69
Ojibwe, 133, 335

Paamese, 72, 92, 94, 97
Permic, 274, 277, 279

Pilagá, 7–13, 15, 149, 150, 150², 150⁴, 150⁵, 151⁶, 151⁷, 152, 153, 153¹⁰, 155, 156, 161, 162, 173, 174, 180
Polish, 25, 56⁶
Purépecha, 335

Quechua, 335

Rama, 335
Russian, 7, 16, 187, 245, 245², 246, 247, 247⁵, 248, 249, 253, 254, 265, 266, 277–279, 286, 286¹⁵, 286¹⁶, 289, 291, 299

Sanskrit, 329
Santali, 322, 335
Semelai, 335
Semitic, 312
Serbo-Croatian, 308, 335
Slave, 335
Slavic, 56⁶, 248
Southeast Ambrym, *see* Vatlongos
Spanish, 25, 27, 28, 37, 132, 160, 168
Supyire, 323, 335
Swedish, 216, 217, 219, 220

Tagalog, 311, 335
Tamashek, 310, 318, 326, 327, 335
Tariana, 335
Tepehua, 335
Tetun, 335
Thai Kamti, 316, 330, 335
Toba, 150⁵, 152, 153, 153¹⁰, 157¹³, 178
Toba Batak, 311, 335
Toqabaqita, 335
Trumai, 335
Tukang Besi, 335
Tümpisa Shoshone, 309, 335
Turkish, 335

Language index

- Tzutujil, 335
- Udmurt, 7, 8, 16, 274, 275, 277, 278,
278⁷, 279, 285–289, 291, 292,
297–300
- Ukrainian, 56⁶
- Ungarinjin, 336
- Urarina, 335
- Urim, 331
- Usan, 298, 299, 315, 316, 331, 332, 336
- Vatlongos, 5, 7, 8, 10–14, 69–75, 75³,
77, 80, 86, 87, 89, 90, 92–94,
97–99
- Vietnamese, 335
- Wambaya, 336
- Wambon, 331–333, 336
- Wan, 9, 14, 44, 45, 45², 46–51, 53–56,
58, 59, 59⁷, 60–64
- Wappo, 329, 330, 335
- Warao, 335
- Wari', 180, 326, 335
- West Greenlandic, 335
- Wichí, 180
- Wolio, 311
- Yagua, 153, 309, 335
- Yimas, 326, 327, 336
- Yukaghir, 335
- Yurakaré, 322, 335

Demonstratives in discourse

This volume explores the use of demonstratives in the structuring and management of discourse, and their role as engagement expressions, from a crosslinguistic perspective. It seeks to establish which types of discourse-related functions are commonly encoded by demonstratives, beyond the well-established reference-tracking and deictic uses, and also investigates which members of demonstrative paradigms typically take on certain functions. Moreover, it looks at the roles of non-deictic demonstratives, that is, members of the paradigm which are dedicated e.g. to contrastive, recognitional, or anaphoric functions and do not express deictic distinctions. Several of the studies also focus on manner demonstratives, which have been little studied from a crosslinguistic perspective. The volume thus broadens the scope of investigation of demonstratives to look at how their core functions interact with a wider range of discourse functions in a number of different languages. The volume covers languages from a range of geographical locations and language families, including Cushitic and Mande languages in Africa, Oceanic and Papuan languages in the Pacific region, Algonquian and Guaykuruan in the Americas, and Germanic, Slavic and Finno-Ugric languages in the Eurasian region. It also includes two papers taking a broader typological approach to specific discourse functions of demonstratives.

ISBN 978-3-96110-286-0



9 783961 102860