The persistence of space: formalizing the polysemy of spatial relations in functional elements

Staps, C.

The persistence of space

Published by

LOT
Binnengasthuisstraat 9
1012 ZA Amsterdam
The Netherlands

phone: +31 20 525 2461

e-mail: lot@uva.nl
https://www.lotschool.nl

Cover illustration: Marga Rabouw

ISBN: 978-94-6093-457-5
DOI: https://dx.medra.org/10.48273/LOT0673
NUR: 616

The persistence of space

Formalizing the polysemy of spatial relations in functional elements

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 20 juni 2024
klokke 16:15 uur

door

Camil Staps

geboren te Leusden
in 1996

Promotores:	Prof. dr. J. E. C. V. Rooryck
	Prof. dr. E. J. van Wolde
Promotiecommissie:	Prof. dr. B. Copley
	Prof. dr. H. Gzella
	Prof. dr. M. G. Kossmann
	Dr. J. Zwarts

The research reported here was supported by a *PhDs in the Humanities* grant awarded by the Dutch Research Council under project number PGW.19.015.

Contents

Acknowledgments

First of all, I would like to thank my PhD supervisors, Johan Rooryck and Ellen van Wolde. Thank you for always making time, answering the most basic of questions, and all your insightful comments on both content and exposition. I could not have hoped for a better supervision team.

The research that led to this dissertation was supported by a *PhDs in the Humanities* grant; I thank the Dutch Research Council for their trust in me. I also thank Holger Gzella, who was indispensable in securing this grant, but also for supervising me during my MA and at the start of my PhD project.

Many people contributed to a stimulating working environment over the past four years; I want to thank in particular Seth Bledsoe, Margaretha Folmer, Jurgen Lingen, Benjamin Suchard, Aren Wilson-Wright, the members of the Biblical Hebrew reading group, and my students.

My friends, not in the least for reminding me of the world beyond these pages: Martijn Beukenhorst; Tom Derie and Lotte Vanfraechem; Ilona Kragtwijk; Katarina Laken; Rinus Plasmeijer; Rick and Marga Rabouw; br. Benoît of Taizé; Erin van der Veen; Xander Vertegaal (also for some LaTeX macros); Bram Visscher (also for clarifying a matter of Belorussian phonology) and Justyna Visscher-Jablonska; Roland van Vulpen; and others.

The seeds for this dissertation had been sown long before I was aware of it. I want to thank my parents Els and Sjef, and my grandparents, for instilling in me a persistent curiosity in language and the world in general, but of course even more for your continuous encouragement. I am also thankful to my brother Merlijn, and Ciara. Finally, my wife Devika, thank you for your endless love and support.

Nijmegen, September 2023

Introduction

In this dissertation I address the question how humans conceptualize abstract relations by analyzing the expression of such relations in natural language. I propose semantic-pragmatic formalisms that capture the kind of polysemy in (1):

(1) a. *Frightened by the guests, the cat fled **out of** the room.*
b. *The cat fled from the room **out of** fear for the guests.*

In (1a), *out of* expresses a spatial relation between *the cat* and *the room*, namely that the cat moved along a path starting in the room and ending outside of it. In (1b), *out of* does not express such a spatial relation. Instead, it expresses a causal relation, whereby the action of *the cat* is the result of *fear for the guests*. The preposition *out of* is polysemous between marking a Source in the spatial domain and marking a Cause in the causal domain. This particular polysemy pattern is very common cross-linguistically, so it is probably not accidental. It is more likely that speakers somehow conceptualize the abstract causal relation with cognitive primitives that are also used to conceptualize spatial relations. I thus argue that aspects of the original spatial meaning do not bleach away but persist in extensions to other domains.

It has long been recognized that many abstract relations are expressed in spatial terms, and within cognitive linguistics this is taken as evidence that spatial representation is somehow primary. However, descriptions of the extension of spatial terms to abstract domains are often given in an intuitionist manner, without explaining how a particular abstract meaning can be *derived* from a spatial meaning. To form a clearer picture of the way spatial representation structures the way we conceptualize abstract relations, we need a constrained theory of the way space is represented and how abstract domains can build on this. This dissertation therefore investigates (a) what kind of spatial notions can be extended to other, abstract domains and (b) what determines their interpretation in these domains. It does so

by offering concrete, formal semantic-pragmatic analyses of these types of polysemy.

In order to obtain a comprehensive picture of the way spatial notions are extended to other domains, this dissertation is not limited to well-known cases of extensions, such as to the temporal domain. Instead, I focus on the following three types of domains: (a) causal relations, (b) relations of speakers to information content, and (c) relations between people. My claim is that all these relations are conceptualized in spatial terms. While I study causal and social relations by means of prepositions, I look at the way information content is positioned in space through the lens of demonstratives and complementizers. The fact that spatial meaning persists in even these highly grammaticalized items provides further evidence for the claim that space is a crucial primitive of human cognition.

Each polysemy pattern is studied in at least two unrelated languages, drawing examples primarily from English, Romance, and Biblical Hebrew. Appendix A provides the reader unfamiliar with Hebrew with the necessary background to work their way through the interlinear examples. Chapters 1 and 3 are primarily based on French and English, respectively.

Of course, the central question of the cognitive conceptualization of abstract relations cannot be answered in the scope of a single dissertation. Nevertheless, I hope that the case studies presented here can convince the reader that this research program has potential, and that the theoretical proposals can serve as a starting point for further investigation.

The chapters of this dissertation are divided into three parts according to the type of abstract relation they discuss. Part I looks at spatial-causal polysemy. Chapter 1 develops a formal model for such polysemy in prepositions, taking Agent phrases in French passives with *de* 'from, by' and *par* 'through, by' as an example. Chapter 2 applies the same model to causal adjuncts introduced by מִן *min* 'from' and בְּ *bə* 'in' in Biblical Hebrew.

Part II is concerned with the use of spatial notions to express relations between speakers and information content. Chapter 3 looks at the English demonstrative and complementizer *that* and explains how the [+distal] feature of the original demonstrative function still plays a role in its use as a complementizer. In chapter 4 the same model is applied to Biblical Hebrew כִּי *kī*, a clausal connective with a plethora of different uses.

Part III, which consists only of chapter 5, considers how spatial terms are used to express relations between people. In particular, it discusses Biblical Hebrew לִפְנֵי *lipnē* 'to the face of' and its English translation 'before', which can both be used to refer to someone's authority or dignity.

Finally, chapter 6 summarizes the main findings.

All main chapters are based on independent articles. As such, there are minor inconsistencies in terminology and assumed background between them. Where needed, I have made slight changes to make the work accessible to a wider audience here. However, I have kept these changes to a minimum so as not to confuse readers with two significantly different versions.

Abbreviations

Glossing abbreviations

1	First person
2	Second person
3	Third person
ALL	Allative
CAUS	Causative Aktionsart (see appendix A.2)
COMP	Complementizer
DU	Dual
F	Feminine
FUT	Future
IMP	Imperative
INF	Infinitive
INFABS	Infinitive absolute, various functions (see appendix A.2)
IPFV	Imperfective (in Biblical Hebrew the prefix conjugation with various meanings; see appendix A.2)
JUSS	Jussive
M	Masculine
MID	Middle voice (niphal template; see appendix A.2)
N	Neuter
NEG	Negation
NOM	Nominative
OBJ	Object marker or pronominal suffix with object function
PASS	Passive voice (see appendix A.2)
PFV	Perfective (in Biblical Hebrew the suffix conjugation with various meanings; see appendix A.2)
PL	Plural
PLURACT	Pluractional Aktionsart (see appendix A.2)
PRS	Present tense

PST	Past tense
PTCP	Participle
Q	Question marker
REFL	Reflexive voice
REL	Relativizer
SG	Singular
WAYQ	*Wayyiqṭol*, narrative sequential simple past (see appendix A.2)
WQAT	*Wəqăṭal*, a sequential modal form (see appendix A.2)

Technical abbreviations

Adj	Adjective
Adv	Adverb
C	Complementizer
CP	Complementizer Phrase
DP	Determiner Phrase
NP	Noun Phrase
PP	Prepositional Phrase
QUD	Question Under Discussion
VoiceP	Voice Phrase
VP	Verb Phrase
*v*P	Little *v* (light verb) Phrase

Abbreviations of Bible translations

ASV	American Standard Version
CSB	Christian Standard Bible
ESV	English Standard Version
KJV	King James Version
NASB	New American Standard Bible
NET	New English Translation
NIV	New International Version
NKJV	New King James Version
NLT	New Living Translation
RSV	Revised Standard Version

Part I

Causation

Formalizing spatial-causal polysemy of Agent prepositions

1

Abstract Current formal approaches to *by*-phrases in passives analyze the Agent preposition *by* as semantically vacuous: the denotation of *by* is merely such that its argument fulfills the same function as the external argument in the corresponding active sentence. This leads to a view of agentive *by* as essentially homonymous with spatial and temporal *by*. We argue, on the basis of work in the cognitive linguistic tradition and a new analysis of the French Agent prepositions *par* and *de*, that Agent markers do have non-trivial semantic content, and are polysemous rather than homonymous with their spatial counterparts. To formalize this we propose to model these prepositions with general schematic denotations of a polymorphic type $\langle \eta, \langle \theta, t \rangle \rangle$, which can be instantiated with a concrete type in a specific syntactic and semantic context, such as $\langle e, \langle e, t \rangle \rangle$ for the spatial meaning of *by*. The use as an Agent preposition is simply one of these instantiations, with type $\langle e, \langle s, t \rangle \rangle$ (where s stands for events). The concrete meaning in context depends on both the general, polymorphically typed denotation and the specific type in the given context. In this way our proposal integrates a useful insight from cognitive linguistics in a semantic formalization of the passive, and opens up possibilities for similar accounts of other highly grammaticalized prepositions.

This chapter was originally published as: Camil Staps & Johan Rooryck. 2024. Formalizing spatial-causal polysemy of Agent prepositions. *Semantics and Pragmatics* 17(4). DOI: 10.3765/sp.17.4. Camil Staps and Johan Rooryck designed the study and collected the data. Camil Staps performed the analysis and wrote the article, in regular discussion with Johan Rooryck.

We thank our informants for verifying our judgments of the French examples in this article. This article has benefited tremendously from insightful comments of Bridget Copley, Louise McNally, Joost Zwarts, the audiences of *Agency and Intentions in Language 2* and the COCOA seminar, and three anonymous reviewers; we extend our gratitude to them as well. Any mistakes are ours alone.

1.1 Introduction

Formal analyses of *by*-phrases in passives tend to treat the Agent preposition *by* as a purely functional element, for example akin to a case marker (Collins 2018; cf. also Bruening 2013). In all accounts known to us, the denotation of *by* is merely what is needed to ensure that its argument plays the same role as the external argument in the corresponding active sentence; *by* does not project any additional meaning. We see three problems with these approaches, which we briefly introduce here.

First of all, approaches that treat the Agent preposition *by* as a purely functional marker effectively take this use of *by* to be accidentally homonymous with its other uses (e.g., spatial *by the house*; temporal *by five o'clock*). They do not clarify the relation between these other uses of *by* and its use as an Agent preposition. This is problematic because the syncretism is not, in fact, accidental; Croft (2012: 222–226) has shown that Agent markers cross-linguistically tend to derive from prepositions with an ablative ('from') or perlative ('through') meaning. For example, English *by* itself had a perlative meaning before obtaining a proximative sense, which survives in expressions like *I went **by** that road* (Palancar 2002: 184). According to Croft, the underlying reason for this would be that causation is cognitively represented as a chain, with causes preceding effects and thus being marked as something 'through' or 'from' which an effect arises. However, if Agent prepositions are purely functional elements, there is a priori no reason why other prepositions, such as *to* or *for*, could not become Agent markers as well. Ideally, the formal analysis would predict that such developments occur only very rarely.

Another argument against these approaches comes from languages that have multiple Agent prepositions, like French.[1] In French passive clauses, the Agent can be introduced by both *par* 'through, by' (1.1a) and *de* 'from, of, by' (1.1b):[2]

1 The French prepositions discussed here have cognates in at least Spanish (Suñer 1981) and Portuguese (Moody 1972: 64–66), with very similar behavior. Our analysis appears to apply to the Agent prepositions in these languages as well, but we focus here on French, as the behavior of the relevant prepositions seems to have been discussed in most detail for this language.

2 Because the difference in acceptability between *de* and *par* can be subtle in many of the examples discussed here, we use superscript numbers to indicate acceptability. These numbers are averaged Likert scores from an informal survey and range from −1 (not

(1.1) a. *Le chien est lavé* $^{1.00}$***par***/$^{-0.96}$***de*** *Marie*
the dog is washed *par*/ *de* Marie

'The dog was washed by Marie.' (Straub 1974: 584)

b. *Le mois de février est précédé* {***du*** /?***par***[3] *le* } *mois*
the month of February is preceded *de*=the/ *par* the month
de janvier
of January

'February is preceded by January.' (Straub 1974: 591)

In some sentences, both prepositions are felicitous, but in different contexts. The choice between *par* and *de* depends on the relation of the Agent to the event. In (1.2), the presence of the parents has little effect when *de* is used; they may simply be observing from off the field. By contrast, when *par* is used, their presence changes the interpretation of the event: they are now more likely to be actively participating in the match.

(1.2) *Les enfants vont jouer au foot accompagnés* $^{0.87}$***de***/$^{0.94}$***par*** *leurs*
the children go play_soccer accompanied *de*/ *par* their
parents
parents

'The children are going to play soccer accompanied by their parents.'
de ⇒ the parents are not very involved; they may be only watching;
par ⇒ the parents may be playing with the children.

Example (1.3) gives an impression of the kind of verbs the French Agent prepositions *de* and *par* are typically used with:

---

acceptable) to 1 (acceptable). They are only meant to give a quick impression of the general tendency of the survey responses; for full details about the distribution of the responses, as well as more information about the survey, see section 1.6. Standard judgment marks are used for sentences that we did not test in our survey.

Note that *de le* shortens to *du*, *de les* shortens to *des*, and *de* reduces to *d'* before vowels. Finally, we are only interested in *de* followed by proper DPs here; for the use of *de* followed by a bare NP see Martin (2005).

3 *Par* is ungrammatical in this example according to Straub (1974), but for speakers we consulted it was acceptable to varying extents in a similar context (see [1.26a] below). Over the years, *par* has become more and more the default preposition. In our tests, speakers only strongly rejected *par* with positional verbs like *précéder* 'precede' and *suivre* 'follow', and then only when the context is clearly stative (cf. [1.1b]). When we critique earlier work, it should be kept in mind that previous analyses may have been correct for older stages of the language, even when they do not apply anymore.

(1.3)		
	par	*briser* 'break'; *construir* 'build'; *écrir* 'write'; *laver* 'wash'; *tuer* 'kill'
	par/*de*	*aimer* 'love'; *respecter* 'respect'; *abandonner* 'abandon'; *délaisser* 'abandon'; *accompagner* 'accompany'; *précéder* 'precede' (dynamic); *suivre* 'follow' (dynamic); *surplomber* 'overlook'[4]
	de	*précéder* 'precede' (stative); *suivre* 'follow' (stative)

We will argue that *de* marks arguments that are less proto-agentive in the event, while *par* marks arguments that are more proto-agentive (in the sense of Dowty 1991 and the scalar notion of transitivity of Hopper & Thompson 1980).[5] This also explains the distribution in (1.1). However, it is difficult to incorporate such information transparently in analyses that treat *by* as a case marker and use the identity function $\lambda x.x$ as its denotation (e.g., Collins 2018), as the event is then not available as an argument of the denotation of the Agent preposition.

A third issue is that current approaches to *by*-phrases are tailor-made for passive sentences (Bruening 2013; Collins 2018; Angelopoulos et al. 2020). However, Agent prepositions are often part of a more general causative pattern of use. For example, *by* can also be used to mark means (*by bus*, *by force*), and French *de* and *par* also have related causal meanings. As we will see below, the denotations proposed for the Agent preposition in accounts of the passive cannot be used in such contexts, because these contexts lack the specific syntactic environment of the passive for which they were developed. The agentive and other causal uses of these prepositions are effectively taken to be accidentally homonymous. However, the semantic contribution of the preposition in these contexts is roughly the same: as we will show below, *de* is associated with stativity in this environment, whereas *par* is associated with dynamicity, a distinction related to proto-agentivity. This suggests that we should be aiming at a more general semantics for these

4 Gaatone (1998: 203) and Spang-Hanssen (1963: 74–76) list many verbs similar to *surplomber* 'overlook': *cerner* 'surround', *encadrer* 'frame', *encercler* 'encircle', *enclore* 'enclose', *entourer* 'surround', *envelopper* 'wrap', *environner* 'surround'; *auréoler* 'halo', *couronner* 'crown'; *border* 'border', *couper* 'cut', (*re*)*couvrir* 'cover', *flanquer* 'flank', *jalonner* 'stake out'.

5 Since these implications survive negation, we assume that this aspect of their meaning is presuppositional. For example, the following is only felicitous if *de* is stressed and used meta-linguistically: *Les enfants ne vont jouer au foot accompagnés* ***de*** *leurs parents; leurs parents participeront aussi* 'The children are not going to play soccer accompanied by their parents; their parents will also participate'.

prepositions, independent of the syntactic structure of the passive. Such an analysis is expected based on Croft (2012), mentioned above, since the notion of a causal chain is not limited to passives but generalizes to other causal contexts.

In sum, we seek an account of *by*-phrases with the following properties:

(1.4) a. The analysis should predict that Agent prepositions only develop from prepositions with specific spatial meanings.

b. The analysis should allow Agent prepositions to express properties of the Agent in relation to the event.

c. The analysis should be general enough to apply to causal uses of Agent prepositions outside passives as well.

We will propose an analysis that satisfies these criteria. Since the prepositions *de* and *par* motivate the need for all three criteria, we specifically focus on French in this chapter. Section 1.2 provides cognitive linguistic background and presents the technical details of the proposal. At the end of this section we show how the account can be generalized from passives to causal adjuncts: it will be useful to discuss *de* and *par* first in this context, before discussing French passives in detail in section 1.3. Section 1.4 discusses related work; descriptive work on French *de* and *par* as well as formal analyses of *by*-phrases and prepositional polysemy. Section 1.5 summarizes and concludes. Section 1.6 describes an informal survey which we used to confirm our judgments for our examples from French.

1.2 Integrating cognitive linguistic insights: the proposal

As already mentioned in the introduction and discussed in more detail in section 1.4.2, formal accounts of *by*-phrases have tended to effectively take the Agent preposition use of *by* as homonymous with the spatial and temporal uses of this preposition. This is in contrast to much work in the cognitive linguistic tradition, which holds that different uses of prepositions are related in a principled way (e.g., Tyler & Evans 2003; Croft 2012):

> [Principled polysemy] holds that a particular form [...] is conventionally associated with a number of distinct but related meanings. [...] In essence [...] our proposal is that (the vast majority of) distinct meaning components associated with a lexical item

> [...] are related to each other in a systematic and motivated way.
> (Tyler & Evans 2003: 37–38)

1.2.1 Incorporating principled polysemy

To incorporate this idea, we propose that these prepositions receive a single denotation, which is general enough to derive, given contextual clues, the different specific meanings of the preposition. Typically, the general meaning has to do with space, since many prepositions can be shown to have developed non-spatial meanings from a spatial origin. In the case of English *by* (e.g. *the house by the lake*), the general meaning would involve *close proximity* of the Figure (*the house*) to the Ground (*the lake*). However, this close proximity is to be understood in an abstract, not necessarily physical way. Thus, for example, moments in time can also be seen as in close proximity to each other.

Our approach makes use of polymorphic types as described by Morrill (1994: 162) (the notion of polymorphic types in Asher 2011: 219–236 is unrelated). The general denotation of the preposition has a polymorphic type $\langle \eta, \langle \theta, t \rangle \rangle$, in which the type variables η and θ can be instantiated with concrete types depending on the syntactic context. In the general schematic denotation in (1.5) we use f and g for Figure and Ground, respectively. The exact formalization of "f is in close proximity to g" depends on assumptions about the cognitive representation of abstract space. For example, Bierwisch (1999: 44) assumes that spatial representation is based on locations in a three-dimensional space, and Zwarts & Winter (2000) develop a more general model based on n-dimensional vectors. In such models, close proximity could be defined in terms of Euclidean distance.

(1.5) $[\![\text{by}]\!]_{\langle \eta, \langle \theta, t \rangle \rangle} = \lambda g_{\eta} \lambda f_{\theta}.f$ is in close proximity to g

To see how the instantiation of a polymorphic type works, consider the physical spatial meaning in *the house by the lake*. For this phrase, the type would be instantiated with $\eta = e$, $\theta = e$, which triggers the specific meaning: applying the abstract notion of close proximity to a context with two concrete entities of type $\langle e \rangle$ gives rise to the specifically physical interpretation of nearness (1.6a). In the case of the causal meaning (*written by Mary*), the preposition describes the relation between a concrete entity and an event,

and must therefore be instantiated with $\eta = e, \theta = s$ (1.6b).[6] We do not take a strong position on what Initiator(x, e) entails exactly. For our purposes this category can be quite broad; for example, Initiator(x, e) could be taken to mean that x is highest on Fillmore's (1968) subject selection scale in e.

(1.6) a. $[\![\text{by}_{\text{spatial}}]\!]_{\langle e,\langle e,t\rangle\rangle} = \lambda x \lambda y.y$ is in close proximity to x
interpretation: y is physically near x

b. $[\![\text{by}_{\text{causal}}]\!]_{\langle e,\langle s,t\rangle\rangle} = \lambda x \lambda e.e$ is in close proximity to x
interpretation: Initiator(x, e)

This approach leads to a clear division of labor. The formal syntactic and semantic context, by using certain types, enforces the required type of the specific instantiation of the polymorphic denotation. This ensures, for example, that we do not interpret *by* spatially in Agent phrases in passives: the context requires an instantiation of type $\langle e, \langle s, t\rangle\rangle$, not $\langle e, \langle e, t\rangle\rangle$. In this way, the formal theory heavily constrains which interpretation a highly polysemous preposition receives.

Without a story about how the different senses of a preposition are related, this formal machinery is still not much more than a way to describe massive homonymy. Using insights from cognitive linguistics we can then explain how the meaning of the concrete instantiations in (1.6) can be derived from the general, schematic denotation in (1.5) and the type provided by the formal context. For the spatial meaning, the explanation will usually be quite simple, since the general denotation is described in abstract spatial terms like "close proximity" in (1.5). For other domains, one needs to consider how that domain is mapped onto the spatial domain or, alternatively, how the spatial representation module is recycled to represent that domain.[7]

There is a long history of research into the spatial representation of causation.[8] Causation is usually considered to build on the spatial notions of

6 This is similar to the "dual analysis" based on reanalysis put forward by Dowty (2003). Formalizing the temporal meaning could be done in a richer type system with separate types for time expressions, but we focus on causal meanings here.

7 For mapping one domain onto another, see Lakoff & Johnson (1980) and subsequent work. For the spatial representation module, see Bierwisch (1999), and for the notion of recycling, see Rooryck (2019) building on Biberauer (2017).

8 The discussion below follows Radden (1985), but similar ideas appear in Dirven (1995) and other sources. Talmy (1988) develops the framework of force dynamics which represents causes as vectorial forces, a theory which we will use in section 1.2.3.

Source and Goal, possibly through an intermediary temporal representation and the common *post hoc ergo propter hoc* fallacy (Radden 1985: 186–194):

(1.7) a. Source, Start, Cause: ***from*** *Paris*, ***from*** *8:30*, *die* ***from*** *hunger*

b. Goal, Endpoint, Purpose: ***to*** *Dijon*, *Monday* ***to*** *Friday*, *dress* ***to*** *impress*

Causation can also build on the notion of Path, in which case multiple interpretations are possible (Radden 1985: 198–200):

(1.8) a. Spatial Path: *pass* ***by*** *a newsstand*

b. Means: *work* ***by*** *candlelight*

c. Permissive Cause: *printed* ***by*** *permission*

d. Agentive Cause: *bitten* ***by*** *a dog*

Croft (2012: 222–226) develops the notion of a causal chain, which contains the different entities that influence each other in an event. A causal chain encompasses the causal Sources, Goals, and Paths. For example, the causal chain for (1.9a) is provided in (1.9b):

(1.9) a. *The coconut was broken for John by Sue with a hammer.*

b. Sue → hammer → coconut → John (cf. Croft 2012: 222)

Based on the mapping of the causal domain onto the spatial domain proposed by Radden (1985), we expect that Source prepositions mark Causes at the origin of the causal chain (e.g., Agents but not Instruments) and that Goal prepositions mark, for instance, Beneficiaries. Path prepositions may be used to mark Means or Instruments, which are between the origin of the causal chain and the Patient. Due to language change, these categories may shift somewhat, so that Path prepositions like *by* commonly mark Agents that appear to be at the origin of the causal chain as well. However, these changes are very limited; for example, we do not expect Path prepositions to mark Beneficiaries. These predictions regarding the causal meaning of spatial prepositions have been confirmed in typological studies (Croft 2012: 225 and references therein). The fact that unrelated languages display the same mappings between spatial and causal concepts suggests a cognitive reality. This forms the basis for a cognitive linguistic argument for the relation between an abstract spatial denotation (1.5) and its causal instantiation (1.6b) (and thereby also the relation with the physical spatial instantiation [1.6a]). We develop this argument further for French in section 1.3.5.

In the following subsections we describe how the proposal in (1.5) can be made to work in passives (section 1.2.2) and illustrate its generality by applying it to prepositions in causal adjuncts (section 1.2.3).

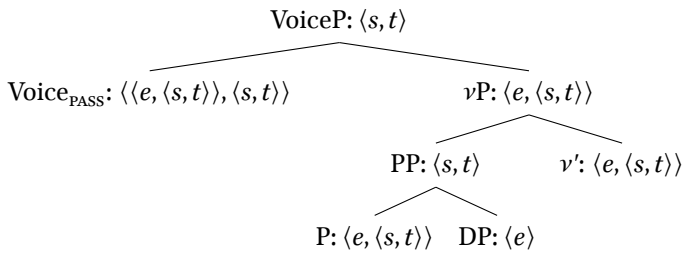
1.2.2 *By*-phrases in passives

To apply the denotation in (1.6b) to passive sentences, we largely adopt the approach to *by*-phrases of Angelopoulos et al. (2020) (see section 1.4.2 for a discussion of the differences, as well as a comparison with other strategies). In this approach, the *by*-phrase takes the same place as the external argument in an active sentence. Given that a *by*-phrase expresses a relation between an Agent and an event, we argue that the denotation of *by* must be of type $\langle e, \langle s, t\rangle\rangle$, following the denotation in (1.6b), repeated here:

(1.6b) $[\![\text{by}_{\text{causal}}]\!]_{\langle e,\langle s,t\rangle\rangle} = \lambda x \lambda e.e$ is in close proximity to x
interpretation: Initiator(x, e)

As a result, the *by*-phrase is of type $\langle s, t\rangle$. In the compositional analysis in (1.10a), we assume a *v*P projection selected by an active or passive Voice head (nothing hinges on this assumption; we make it to simplify our comparison with other accounts in section 1.4.2). The *by*-phrase combines with the v' projection using Event Identification (Kratzer 1996: 122). It fills the syntactic position of the external argument, thus preventing a *by*-phrase from occurring in active sentences, without saturating the semantic argument. All that is left for the Voice head is to perform existential closure (1.10b), which is redundant in the case of a passive with *by*-phrase but necessary in a passive without *by*-phrase.

(1.10) a.



b. $[\![\text{Voice}_{\text{PASS}}]\!] = \lambda p \lambda e.\exists x\colon p(x)(e)$

Example (1.11) provides an example of the derivation of a passive with a *by*-phrase. We gloss over the derivation of the v' projection here. The existential closure introduced by $\text{Voice}_{\text{PASS}}$ is needed in passives without *by*-

phrases, but becomes redundant in this derivation when the variable it introduces is identified with the argument of *by*. This can be ensured via a principle such as Chomsky's (1981) theta criterion: an event can only have one Initiator (Landman 2000: 68; Williams 2015: 287; cf. Dowty 1989: 85, 99–103).

(1.11) a. $[\![v'$ *the book was sent*$]\!] = \lambda x \lambda e.\mathbf{send}(e)$ & Initiator(x, e)
& Patient$(\iota y.\mathbf{book}(y), e)$

b. $[\![$*by John*$]\!] = \lambda e.$ Initiator$(\iota y.\mathbf{John}(y), e)$

c. $[\![v\mathrm{P}]\!] = \lambda x \lambda e.\mathbf{send}(e)$ & Initiator(x, e) & Patient$(\iota y.\mathbf{book}(y), e)$
& Initiator$(\iota y.\mathbf{John}(y), e)$ (Event Identification)

d. $[\![\mathrm{VoiceP}]\!]$
$= \lambda e.\exists x\colon \mathbf{send}(e)$ & Initiator(x, e) & Patient$(\iota y.\mathbf{book}(y), e)$
& Initiator$(\iota y.\mathbf{John}(y), e)$
$= \lambda e.\mathbf{send}(e)$ & Patient$(\iota y.\mathbf{book}(y), e)$ & Initiator$(\iota y.\mathbf{John}(y), e)$
(theta criterion)

The analysis in (1.10) places *by* on the border between a purely lexical and a purely functional preposition. On the one hand, its type is regular and it carries semantic content related to the general meaning in (1.5); on the other hand, the *by*-phrase appears in the same syntactic specifier position as the external argument in active sentences. This intermediate status is in line with the fact that *by* has both lexical (e.g., spatial) and functional (e.g., agentive) uses.

1.2.3 Prepositions in causal adjuncts

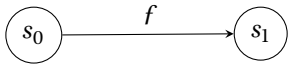
Before turning to the interpretation of French *de* and *par* in passives, we illustrate the generality of our proposal by showing how it can be used to capture distributional facts about prepositions in causal adjuncts. The example we work out here is the observation by Copley & Harley (2015) that English *from* marks causes that are forces, rather than causes that are situations:

(1.12) a. *The floor broke* ***from*** *the *(weight of the) elephant.*
(Copley & Harley 2015: 141)

b. *The window broke* ***from*** *John*('s hitting it).*
(Copley & Harley 2015: 141)

Based on this and many other facts, Copley & Harley (2015, 2022) develop a semantic framework with primitive types for situations (type $\langle s \rangle$)

and forces (type $\langle f \rangle$), rather than events (in this subsection we thus use s for situations and not events). Conceptually, a situation "includes individuals and their property attributions" (Copley & Harley 2022: 12; cf. Barwise & Perry 1983), and a force is an input of energy that arises from a situation. Formally, a force is a function that maps situations to situations: $f(s_0) = s_1$. In this framework, a causal event is not a composite of a cause and a result, but a force that maps one situation to another (1.13). The distribution of *from* can now be captured in terms of a type constraint: causal *from* has type $\langle f, \langle s, t \rangle \rangle$ but not $\langle e, \langle s, t \rangle \rangle$ or $\langle s, \langle s, t \rangle \rangle$.

(1.13) 

(Copley & Harley 2022: 12)

French *de* and *par* are sensitive to the same distinction between situations and forces. The examples in (1.14) show that *de* can be used to name causes that are situations (e.g., *faim* 'hunger'), while it cannot be used to name causes that are forces (e.g., *un tremblement de terre* 'an earthquake'). Instead, a majority of speakers use *par* for this purpose:

(1.14) a. *Jean est mort* $^{1.00}$***de***/$^{-0.87}$***par*** {*faim* / *vieillesse* / *la maladie de Parkinson*}
Jean is dead *de*/ *par* hunger / old_age / the disease of Parkinson
'Jean died of/from hunger/old age/Parkinson's disease.'

b. *La fenêtre s'est cassée* $^{-0.90}$***de***/$^{0.37}$***par***[9] {*un tremblement de terre* / *l'impact du ballon*}
the window REFL=is broken *de*/ *par* an earthquake / the=impact of=the ball
'The window broke due to an earthquake/the impact of the ball.'

We therefore propose that *de* is instantiated with type $\langle s, \langle s, t \rangle \rangle$, but *par*, for most speakers, with type $\langle f, \langle s, t \rangle \rangle$.[10] This can be formalized with the

9 There is much inter-speaker variation in the acceptability of *par* here. One speaker for whom *par* was unacceptable suggested that *par* is better in sentences like *Ce récipient peut se fermer par une simple pression du couvercle* 'This recipient can be closed by simply pressing on the lid', which are more focused on the process than the result.

10 Thus the type of English *from* matches that of *par*, and not that of *de*, as we might expect based on spatial meaning. This could be because English partitions the causal space differently with a contrast between *from* and *of*, while the immediate parallel for French *par*, English *through*, is not as frequent in causal adjuncts.

concrete denotations in (1.15b; 1.16b), based on the general, schematic denotations in (1.15a; 1.16a). Here, net(s) is the net force generated in situation s (Copley & Harley 2022: 14). The approach is analogous to that for *by* in (1.5–1.6), save for the use of a different type system.

(1.15) a. $[\![\text{par}]\!]_{\langle \eta,\langle \theta,t\rangle\rangle} = \lambda g_{\eta}\lambda f_{\theta}$.figure f is through/via ground g

b. $[\![\text{par}_{\text{causal}}]\!]_{\langle f,\langle s,t\rangle\rangle} = \lambda f\lambda s$.situation s is through/via force f
interpretation: s comes about through f
formally: $\exists s_0$: net$(s_0) = f$ & $f(s_0) = s$[11]

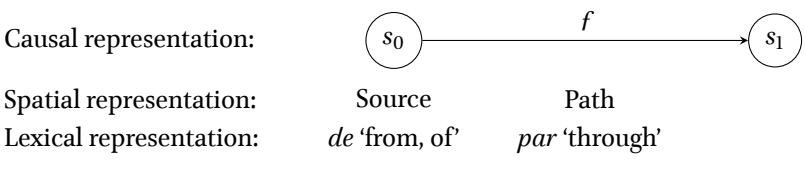
(1.16) a. $[\![\text{de}]\!]_{\langle \eta,\langle \theta,t\rangle\rangle} = \lambda g_{\eta}\lambda f_{\theta}$.figure f is from/of ground g

b. $[\![\text{de}_{\text{causal}}]\!]_{\langle s,\langle s,t\rangle\rangle} = \lambda s\lambda s'$.situation s' is from/of situation s
interpretation: s' arises from s
formally: $(\text{net}(s))(s) = s'$

In words, *par f* expresses that f is the net force of a situation s_0, and that f maps s_0 to the situation s described in the clause. Thus, in (1.14b), the earthquake (or the impact of the ball) is the net force f of a situation s_0, so that $f(s_0) = s$ is a situation in which the window is broken. By contrast, *de s* expresses that the net force of s maps s to the situation s' described in the clause. In (1.14a), s contains Jean, who suffers from hunger (etc.), and is such that its net force brings about s' in which Jean has died.

To be sure, both sentences with *de* and sentences with *par* represent the causal event as in (1.13). However, *par* names the force, and *de* the causing situation. This is not accidental. Recall from section 1.2.1 that when a preposition develops a causal meaning, the position in the causal chain marked by that preposition depends on its spatial meaning (Croft 2012: 222–226). Similarly, we can see (1.13) as a spatial representation of a causal event. In this representation, s_0 can be seen as a Source, and f as a Path. The choice of preposition for each argument is based on its spatial meaning:

(1.17)



11 Copley & Harley (2015: 142) give *from* the denotation $\lambda f\lambda s$. net(pred(s)) = f, with pred(s) defined as the predecessor situation of s. This denotation is roughly the same as the one in (1.15b), but we do not assume that a situation's predecessor is identifiable.

In this way, the conceptualization of causation using forces provides a cognitive linguistic account for the derived meanings in (1.15–1.16). As such, it accounts for the fact that *de* marks situations and *par* marks forces.[12]

1.3 Polymorphically typed semantics for *de* and *par* in passives

In this section we extend the analysis of section 1.2.3 to the agentive meanings of *de* and *par*. The main part of this section consists of a description of the distribution of *de* and *par* in passives. As already mentioned in the introduction, *de* is preferred for less proto-agentive arguments, whereas *par* is used for more proto-agentive arguments. Note that this distribution fits to the use of *de* to mark situations and the use of *par* to mark forces in causal adjuncts (section 1.2.3): forces are associated with dynamic events, which have higher transitivity according to Hopper & Thompson (1980). In sections 1.3.1 to 1.3.4 we show that various aspects of proto-agentivity play a role in the choice between *de* and *par* in passives. Section 1.3.5 then shows how these facts can be accounted for in the analysis proposed in section 1.2.2, and address the question why *de* and *par* might be sensitive to proto-agentivity.

The factors we found to be relevant for the choice between *de* and *par* are a combination of proto-Agent properties (Dowty 1991: 572) and proto-transitivity properties (Hopper & Thompson 1980).[13] All relevant factors are relational properties in the sense of Næss (2007: 30–32); they concern the relation of the Agent to the event. They are the stative/dynamic contrast ("kinesis" in Hopper & Thompson 1980), telicity ("punctuality"), volitional-

12 As pointed out to us by Louise McNally (p.c., October 20, 2023), it is also possible to account for the distribution of *de* and *par* by giving *de* a highly underspecified meaning like that of English *of* (e.g. Partee 1997). *De* could then be excluded from marking forces because there already is a dedicated preposition for forces (*par*). We prefer the semantics in (1.16) because French *de* is much more clearly spatial than English *of*, also covering the meaning of *from*. Furthermore, *de* is not unmarked: in passives, its use is highly restricted, and it is actually *par* which is used as the default Agent preposition.

13 There are correlations with proto-Patient properties and affectedness (Beavers 2011), but these are indirect. *Par* will be used more with highly affected Patients, but only because *par* is used to mark Agents that bring about a change, which go together with highly affected Patients. To see that *par* does not express affectedness of the Patient, consider that many verbs that imply a little affected Patient in the hierarchy of Beavers (2011: 358) take *par*, such as *voir* 'see', *considérer* 'consider', and *lorgner* 'ogle'.

ity, and bringing about a change ("agency", "potency").[14] Of these, the property of bringing about a change is primary, in the sense that if a verb can imply a change, the use of *par* will force it to do so.

We will not address the question where the threshold of "high" and "low" proto-agentivity lies, exactly. In intermediate cases, where the Agent has some but not all properties of proto-Agents, it is to be expected that speakers show quite some variation as to their preference for one preposition or the other. This should be the topic of a more descriptively oriented study. Among our survey participants we could not clearly distinguish clusters of speakers with similar preferences. Here we only aim to (a) establish the fact that there is a proto-agentivity threshold that determines the choice between the two prepositions, and (b) propose a theory to account for it.

1.3.1 Change: prototypically transitive verbs

Prototypically transitive verbs by definition take an Agent that is high in proto-agentivity. In this subsection we treat verbs that imply at least that the Agent brings about a change (whether physical or not). We use Beavers's (2011) conception of *affectedness* to define change. For Beavers, affectedness involves (a) a Theme participant undergoing a change and (b) a scale participant measuring the change.[15] Since we are dealing with passive sentences, we use the term "Patient" rather than "Theme", except when discussing Incremental and Holistic Themes.

In prototypically transitive events, the Agent volitionally and telically causes a physical change in a Patient, as in (1.1a), repeated below. In this example the Patient/Theme is the dog and the scale is being-washed or cleanliness:

(1.1a) *Le chien est lavé* $^{1.00}$***par***$/^{-0.96}$***de*** *Marie*
the dog is washed *par*/ *de* Marie
'The dog was washed by Marie.' (Straub 1974: 584)

In (1.3) above we gave more examples of highly transitive verbs which only take *par*. We also consider verbs of maintaining to belong to this group:

14 As an anonymous reviewer points out, based on these properties we may expect the degree of transitivity to depend on aspect. As a result, some of the judgments we provide in this chapter may be different if the aspect of the sentence is changed.

15 However, we cannot depend on Beavers (2011) too directly, as he explicitly limits himself to dynamic predicates, while many of our examples involve stative predicates.

(1.18) *Le bord supérieur du filet est maintenu* {$^{1.00}$***par*** *des* / $^{-0.81}$***des*** } *flotteurs et demeure à la surface*
the edge upper of=the net is maintained *par* of=the/ *de*=the floats and remains on the surface

'The upper edge of the net is buoyed with floats and remains on the surface.'[16]

This is a case of entrainment causation (cf. Michotte 1946 in citation by Copley & Harley 2022: 4–5; see also the discussion of "maintenance" by Neeleman & van de Koot 2012: 38–43 and "stative causers" by Kratzer 2000 and Pylkkänen 1999). In entrainment causation, the effect occurs during the cause rather than after the cause (which is *launching causation*). For example, in *push the cup **to** the edge of the table*, the cup is at the edge *after* the pushing (launching causation), but in *push the cup **along** the edge of the table*, the cup is along the edge *during* the pushing (entrainment causation). In the latter case, there is no change in the along-the-edge-ness of the cup, which is nevertheless brought about by the pushing. Entrainment causation thus provides a middle ground between a lack of causation (in which no participant is causally affected) and launching causation (in which a change in the described state can be observed). Similarly, in (1.18) there is no physical change, but there is physical causation. The scale measures whether the net is on the surface (or, alternatively, the depth of the net), and the Agent is needed to keep the Patient at the same position on that scale. There is also volitionality, since the floats are placed purposefully. These features entail relatively high proto-agentivity, which explains the preference for *par*.

Verbs with Incremental and Holistic Themes (Dowty 1991: 567–571) also belong to this group. With an Incremental Theme, the scale to measure change is derived from the extent of the Theme (1.19). Clearly, the Agent is highly proto-agentive due to the clear change it brings about in the Theme.

(1.19) *Le gâteau a été mangé / cuit* ***par***/***de*** *Jean*[17]
the cake has been eaten / baked *par*/ *de* Jean

'The cake was eaten/baked by Jean.'

With a Holistic Theme, the Theme is conceived of as a path that maps onto a scale on which change can be measured. Thus, in (1.20a), the degree

16 http://tsb.gc.ca/fra/rapports-reports/marine/2004/m04w0225/m04w0225.html, retrieved January 30, 2023.

17 This example was not included in our survey, but is uncontroversial.

to which the route has been followed is measured by the point on the route, and similarly for (1.20b).

(1.20) a. *Voici la route suivie* {$^{1.00}$***par*** *les* /$^{-0.45}$***des*** } *premiers*
see=here the route followed *par* the/ *de*=the first
explorateurs qui sont arrivés en Amérique
explorers who are arrived in America
'This is the route followed by the first explorers who arrived in America.' (based on Gaatone 1998: 203)

b. *Le désert était traversé* $^{0.96}$***par***/$^{-0.77}$***de*** *la caravane*
the desert was crossed *par*/ *de* the caravan
'The desert was crossed by the caravan.'[18]

While one could argue that the route in (1.20a) only comes into existence in and because of the described event, the desert in (1.20b) cannot be said to be brought about or affected by the caravan. In this case the change is not in the Patient but in the Agent itself. In this sense the Agent is still involved in bringing about a change, namely in its own position (also cf. Dowty's 1991: 572 proto-Agent property "movement (relative to the position of another participant)").

In sum, while the exact cut-off point will vary between speakers, it is clear that there is a group of highly transitive verbs that require *par*. This group contains at least telic verbs that entail physical change, verbs of maintaining, and verbs with Incremental/Holistic Themes.

1.3.2 Change on a contextually inferred scale

With some verbs that do not imply a change in and of themselves, change can be implied by the use of *par* when a scale can be inferred based on the context. We are only aware of examples of stative verbs, so all the examples in this subsection are cases of entrainment causation.

Inferred scales are particularly frequent with emotion verbs. Being stative, emotion verbs have been reported as preferring or requiring *de* (Clédat 1900 and, to a lesser extent, Straub 1974), but we now see that *par* is available with these verbs as well and appears to be taking over as the default.

18 *Traverser* 'traverse' also occurs with *de*, but then selects a bare NP without article: *un espace traversé de/*des tensions politiques* 'a field riddled with political tensions'. This is a genitive of substance (Martin 2005) and is unrelated.

Nevertheless, *de* remains quite acceptable for most speakers in our survey. It is now used in particular when the emotion is presented as not having any effect. Thus, in (1.21a), the love of the grandfather has no effect beyond his own emotional state. By contrast, in (1.21b), the love of the grandfather is the cause of concrete actions, which affect the Patient:[19]

(1.21) a. *Elle est adorée* $^{0.60}$***de***/$^{0.77}$***par*** *son grand-père qui devient*
she is loved *de*/ *par* her grandfather who becomes
toujours émotionnel quand il regarde ses photos
always emotional when he looks_at her photos
'She is loved by her grandfather, who always gets emotional when he looks at her photos.'

b. *Elle est adorée* $^{0.60}$***de***/$^{0.77}$***par*** *son grand-père qui l'emmène*
she is loved *de*/ *par* her grandfather who her=takes
toujours manger des glaces et lui offre d'énormes
always eat of=the ice_creams and her gives of=huge
cadeaux pour son anniversaire
presents for her birthday
'She is loved by her grandfather, who always takes her to eat ice cream and gives her huge presents for her birthday.'

The mention of concrete actions on the part of the Agent (here Experiencer) in (1.21b) suggests that the *adorer* event implies a change on a being-spoiled scale. No scale for change can be inferred in (1.21a). The lower degree of proto-agentivity in (1.21a) compared to (1.21b) explains why *de* is more, and *par* less acceptable in (1.21a) than (1.21b). When the context is not rich enough, either preposition will be felicitous, but the use of *de* will suggest that the event is relatively inconsequential.[20]

19 One may compare *He sneezed the napkin off the table*, where *sneeze* atypically brings about a change on a contextually inferred location scale (Beavers 2011: 360; Boas 2003: 260–277; also cf. Bar-Asher Siegal & Boneh 2020: 38–43).

20 The choice of Agent preposition with emotion verbs has received quite some attention in the literature on Romance languages. Moody (1972: 66) suggests that the loving is Platonic in Portuguese *Nora é amada* ***de*** *todos* 'Nora is loved by all', but that with *por* (French *par*) "an entirely different event may be implied". For Clédat (1900: 222–223), *adoré par* is only felicitous in the sense of 'worship' (*Les animaus* [sic] *sont adorés* ***par*** *certains peuples* 'Animals are worshiped by certain nations'), which may imply consequences such as offerings or vegetarianism. According to him, *de* is required in both contexts in (1.21). This must reflect an older stage of the language, however, since Straub

The judgments for (1.22–1.23) with *aimer* 'love' and *respecter* 'respect' are similar, but the difference is not as large:

(1.22) a. *Le prêtre était très aimé* $^{0.92}$***de***/$^{0.70}$***par*** *ses paroissiens*
the priest was very loved *de*/ *par* his parishioners
parce qu'il était toujours attentif à leurs besoins
because=he was always attentive to their needs
'The priest was much loved by his parishioners because he was always attentive to their needs.'

b. *Le prêtre était très aimé* $^{0.83}$***de***/$^{0.77}$***par*** *ses paroissiens; ils*
the priest was very loved *de*/ *par* his parishioners; they
lui donnaient toujours des tartes et des bouteilles
him gave always of=the cakes and of=the bottles
de vin
of wine
'The priest was much loved by his parishioners; they always gave him cakes and bottles of wine.'

(1.23) a. *Il était évident qu'il s'agissait d'un roi très respecté*
it was evident that=it REFL=dealt of=a king very respected
$^{0.77}$***de***/$^{0.89}$***par*** *sa communauté et* ***de***/***par*** *la société dans*
de/ *par* his community and *de*/*par* the society in
son ensemble
its whole
'It was clear that this was a king who was much respected by his community and the society as a whole.'

b. *Le roi était très respecté* $^{0.64}$***de***/$^{0.81}$***par*** *ses sujets qui*
the king was very respected *de*/ *par* his subjects who
lui apportaient du tribut chaque année
him brought of=the tribute every year
'The king was much respected by his subjects who brought him tribute every year.'

We hypothesize that this has to do with the greater stativity of these verbs. For example, *aimer* and *respecter* combine well with *ne ... plus* 'no longer', while *adorer* does not (e.g., *il* {*n'aime*/*ne respecte*/??*n'adore*} *plus le*

(1974: 586) already reported that *Le garçon est adoré* ***par*** *le grand-père* 'The boy is loved by his grandfather' is felicitous.

professeur 'he does not love/respect/worship the professor anymore'; see Katz 2003 for this test in English).

There are more types of verbs that can imply change on a contextually inferred scale. We already discussed (1.2), repeated here, in section 1.1. When the parents are involved in the event, they are marked by *par*; *de* is preferred when they are less involved, for example when they are merely watching. When the parents join in the event with *par*, this does not necessarily imply a change, but it might: the game may get rougher, for example. The potential for change is already enough to trigger the use of *par*, since *de* would imply that the accompaniment by the parents has no effect at all.[21]

(1.2) *Les enfants vont jouer au foot accompagnés* $^{0.87}$***de***/$^{0.94}$***par*** *leurs parents*
the children go play_soccer accompanied *de*/ *par* their parents

'The children are going to play soccer accompanied by their parents.'
de ⇒ the parents are not very involved; they may be only watching;
par ⇒ the parents may be playing with the children.

A minimal pair can be constructed as in (1.24).[22] In (1.24a), the policeman is guarding the inmate, which is seen as a form of (non-physical) affecting. We can understand this in two ways. Either the policeman psychologically affects the inmate, or there is a potential for change: if the inmate tries to escape, the policeman will try to prevent this. By contrast, (1.24b) involves a *former* inmate who merely happens to be accompanied by a policeman. Most of our informants found both *de* and *par* acceptable in both sentences, but several commented that *par* foregrounds the aspect of surveillance.[23]

21 Rappaport Hovav & Levin (2001: 787–788) and Beavers (2011: 357–365) discuss potential change in slightly different contexts, lending support to the idea that potential for change is positively correlated with transitivity.

22 We thank an anonymous reviewer for suggesting this contrast.

23 The difference may be brought out better if the context in (1.24a) were such that the inmate is more likely to escape (and therefore needs surveillance): *Le détenu se rend aux funérailles de sa mère pendant sa liberté conditionnelle, accompagné* ***par***/***#d'*** *un policier* 'The prisoner is going to his mother's funeral during his parole, accompanied by a policeman.' A similar contrast is discussed by Moody (1972: 66) for Portuguese: *O presidente fugiu seguido* ***da***/***pela*** *polícia* 'The president fled followed by the police.' When *da* is used, the police "did not act upon the president" (e.g., after a coup the president is followed by the police forces loyal to him); but with *pela* (French *par*), the police "pursued" the president (e.g., after the president has escaped with the country's treasure).

(1.24) a. *Le détenu se rend au poste médical accompagné*
the prisoner REFL goes to=the station medical accompanied
$^{0.96}$***par***/$^{0.79}$***d'*** *un policier*
par/ *de* a policeman
'The prisoner is going to medical accompanied by a policeman.'

b. *L'ex-détenu est apparu devant le tribunal,*
the=ex=prisoner is appeared before the courthouse
accompagné {$^{0.89}$***par*** *le* /$^{0.79}$***du*** } *policier qui*
accompanied *par* the/ *de*=the policeman who
l'avait arrêté
him=had arrested
'The former prisoner appeared in front of the courthouse accompanied by the policeman who had arrested him.'

For some speakers, the possibility of implying a change on a contextually inferred scale is not limited to animate Agents. In (1.25b), the inanimate mountain chain keeps the value of the village on the scale measuring the speed with which the emergency services arrive below a certain threshold. By contrast, there is no such scale in (1.25a), where any effect of the surrounding mountains is explicitly denied. Thus, (1.25b) implies a change on a contextually inferred scale (the lateness of the emergency services).

(1.25) a. *Le village est entouré* $^{0.87}$***d'*** /$^{0.73}$***par*** *une chaîne de*
the village is surrounded *de*/ *par* a chain of
montagnes, mais néanmoins bien relié au reste
mountains but nevertheless well connected to=the rest
du pays
of=the country
'The village is surrounded by a mountain chain, but nevertheless well-connected to the rest of the country.'

b. *Le village est entouré* $^{0.79}$***d'*** /$^{0.77}$***par*** *une chaîne de*
the village is surrounded *de*/ *par* a chain of
montagnes, à cause de laquelle les services d'urgence
mountains at cause of which the services of=emergency
arrivent toujours trop tard
arrive always too late
'The village is surrounded by a mountain chain, because of which the emergency services always arrive too late.'

Our judgments for this pair were confirmed by only a few survey respondents; for most, *de* and *par* were equally acceptable. This may be because the mountain chain in (1.25), as a non-animate Agent, does not have volitionality, while differences in volitionality could be a contributing factor to the choice of Agent preposition for the previous examples in this section.

1.3.3 Volitionality: *suivre* 'follow'

There are also verbs for which the difference in interpretation expressed by *de* and *par* does not involve change but volitionality. This is most clear with verbs like *suivre* 'follow' and *précéder* 'precede', that have both a dynamic and a stative, generic reading.[24] The clearest difference exists between a purely locative and a goal-oriented, volitional interpretation. The use of *par* is not quite acceptable for all speakers in the former case (1.26a), while it is required for the standard reading of (1.26b).[25]

(1.26) a. *Lundi est précédé* $^{0.83}$***de***/$^{0.50}$***par*** *dimanche et le mois*
Monday is preceded *de*/ *par* Sunday and the month
de février est précédé {$^{0.83}$***du*** /$^{0.50}$***par*** *le* } *mois de*
of February is preceded *de*=the/ *par* the month of
janvier
January
'Monday is preceded by Sunday and February is preceded by January.'

b. *Le criminel est suivi* {$^{0.35}$***du*** /$^{0.98}$***par*** *le* } *detective*
the criminal is followed *de*=the/ *par* the detective
qui voulait le prendre en flagrant délit
who wanted him take red-handed
'The criminal is followed by the detective who wanted to catch him red-handed.'

Example (1.26a), together with minimal pairs like (1.27), suggests that *de* is used more in generic contexts. We share this intuition, but we believe this

24 For *suivre* 'follow a path' rather than 'follow something/someone', see (1.20a) above.

25 The use of *de* suggests a purely spatial relation between the criminal and the detective (thus decreasing the Agent's proto-agentivity). Such a reading was meant to be excluded by *qui voulait le prendre en flagrant délit* 'who wanted to catch him red-handed', but, judging from survey comments, some informants marked *de* as acceptable here because the relation may still be purely spatial, for example if the detective is unknowingly, accidentally following the criminal.

to be a side effect of properties of proto-agentivity. Generic statements can be used to mention things that depend on convention (1.26a; 1.27a), whereas a concrete statement like (1.27b) more often involves volitionality; in this case the author's volitional choice to order the chapters in this way.

(1.27) a. *Le dernier chapitre est suivi* $^{0.98}$***d'*** /$^{0.31}$***par*** *une table*
the last chapter is followed *de*/ *par* a table
des matières
of=the contents
'The last chapter is followed by a table of contents.'

b. *Cette introduction est suivi* $^{0.77}$***de***/$^{0.64}$***par*** *l'étude de la*
this introduction is followed *de*/ *par* the=study of the
structure atomique et électronique des atomes
structure atomic and electronic of=the atoms
'This introduction is followed by the study of the atomic and electronic structure of atoms.'[26]

To show that the generic flavor of (1.26a; 1.27a) is only a side effect, consider the pair in (1.28). These sentences are equally generic, yet *de* is clearly preferred in (1.28a), while the difference is smaller in (1.28b).

(1.28) a. *Pour mettre en place l'échiquier, on place les pions*
to set in place the=chess_board one places the pawns
sur la deuxième rangée, suivis {$^{0.89}$***des*** /$^{0.50}$***par*** *les* }
on the second rank followed *de*=the/ *par* the
autres pièces sur la première rangée
other pieces on the first rank
'To set up the chess board, we place the pawns on the second rank, followed by the other pieces on the first rank.'

b. *Dans l'ouverture, nous avançons d'abord quelques pions,*
in the=opening_game we advance at_first some pawns
suivis {$^{0.77}$***des*** /$^{0.68}$***par*** *les* } *cavaliers*
followed *de*=the/ *par* the knights
'In the opening game, we first advance some pawns, followed by the knights.'

26 Based on https://www.programmes.uliege.be/cocoon/20212022/cours/CHIM9275-1.html, retrieved December 8, 2021.

The difference between these sentences lies in the volitionality of the presupposed chess player. In (1.28a), there is no strong reason to set up the pawns first. It may be slightly more practical (setting up the other pieces first would require lifting the pawns over the other pieces to place them on the second rank), but nothing would go wrong if one were to set up the pieces in a different order instead, for example from left to right. In (1.28b) however, the player has good reason to advance the pawns first: they can be used to control the center, while at the same time preparing the queen and bishops for development. There is clear purpose behind the decision to advance the pawns first, even though this purpose (and hence volitionality) is ascribed to the presupposed player rather than the pieces themselves.

The same type of volitionality, and hence proto-agentivity, also explains the preference for *par* in (1.29b): sending out the infantry before the cavalry is part of a well-thought-out strategy. By contrast, the order in (1.29a) is determined by protocol, and involves less purpose and volitionality.

(1.29) a. *Au défilé militaire du 14 juillet, l'infanterie était*
at=the parade military of=the 14 July the=infantry was
suivie $^{0.71}$***par***/$^{0.81}$***de*** *la cavalerie*
followed *par*/ *de* the cavalry

'In the military parade of July 14, the infantry was followed by the cavalry.'

b. *Napoléon envoya l'infanterie au combat, suivie*
Napoleon sent_out the=infantry to=the battle followed
$^{0.77}$***par***/$^{0.73}$***de*** *la cavalerie*
par/ *de* the cavalry

'Napoleon sent out the infantry to battle, followed by the cavalry.'

Example (1.30) presents an interesting case:

(1.30) *Ce pianiste est toujours suivi* $^{0.98}$***par***/$^{0.66}$***d'*** *une foule*
this pianist is always followed *par*/ *de*=a crowd
d'admirateurs
of=admirers

'This pianist is always followed by a crowd of admirers.'
de ⇒ the admirers are physically behind the pianist;
par ⇒ the admirers could also be following the pianist's career.
(based on Gaatone 1998: 203)

Both *de* and *par* are felicitous here, but *de* suggests a spatial relation, whereas *par* suggests that the admirers are following the pianist's *career*. There are no obvious differences in the properties of verbal aspect, volitionality, telicity, or bringing about a change. It may be that the difference in interpretation is simply due to the frequent use of *de* with *suivre* in purely locative contexts similar to (1.26a), but this explanation is ad hoc. It rather seems to us that there is a subtle difference in volitionality. With *de*, it is likely that the crowd does not consist of the same members in each instance of the habitual event: if the pianist is on tour, the crowd will likely be different in each city. A reading in which the members of the crowd change regularly is much less likely with *par*, it seems to us. Here we understand a dedicated group of admirers that persistently follows the pianist's career. This dedication could be understood as relating to a higher degree of volitionality, and hence proto-agentivity. However, it is clear that more minimal pairs with better contexts would have to be tested to verify this.

1.3.4 Telicity: *abandonner* 'abandon' and *délaisser* 'leave behind'

Finally, with some verbs, the choice between *de* and *par* tells us something about telicity. It has long been recognized that *de* is not always permitted when a goal PP is added, making the event telic:[27]

(1.31) a. *Un enfant abandonné* ***de***/***par*** *ses parents*
a child abandoned *de*/*par* its parents
'a child abandoned by its parents' (Authier-Revuz 1972: 50)

b. *Un enfant abandonné* ****de***/***par*** *ses parents sous le porche*
a child abandoned *de*/*par* its parents under the porch
'a child abandoned by its parents under the porch'
(Authier-Revuz 1972: 50)

In (1.31b), there has clearly been an event of physically abandoning the child, whereas (1.31a) could be used for a neglected child (and that is certainly the interpretation triggered by *de*). We attempted to capture this contrast in a minimal pair with a difference between children needing food and accommodation due to their parents' abandonment (telic; [1.32a]) and children needing help with homework and social problems (atelic; [1.32b]). An-

27 These examples were not included in our survey, but are uncontroversial.

other minimal pair tested a similar opposition with *délaisser* 'abandon, neglect': (1.33a) is telic; (1.33b) atelic (*laisser* 'leave' behaves the same).

(1.32) a. *Notre organisation soutient les enfants abandonnés* $^{-0.12}$***de***/
our organization helps the children abandoned *de*/
$^{1.00}$***par*** *leurs parents avec un logement et de la*
par their parents with an accommodation and of the
nourriture
food
'Our organization helps children abandoned by their parents with housing and food.'

b. *Notre organisation vient en aide aux enfants abandonnés*
our organization comes in aid to=the children abandoned
$^{-0.07}$***de***/$^{0.92}$***par*** *leurs parents, et les aide à faire leurs*
de/ *par* their parents and them helps at doing their
devoirs et résoudre leurs problèmes sociaux
homework and resolving their problems social
'Our organization comes to the aid of children abandoned by their parents and helps them with doing their homework and resolving social problems.'

(1.33) a. *Quand l'alarme a sonné, Notre Dame a été vite*
when the=alarm has sounded Notre Dame has been quickly
délaissée {$^{0.10}$***des*** /$^{0.83}$***par*** *les* } *touristes qui s'y*
abandoned *de*=the/ *par* the tourists who REFL=there
trouvaient
found
'When the alarm rang, the Notre Dame was quickly abandoned by the tourists who were there.'

b. *Voici une photo de Notre-Dame délaissé* [sic] $^{0.28}$***de***/$^{0.85}$***par***
see=this a photo of Notre Dame abandoned *de*/ *par*
ses touristes en plein confinement pendant COVID
its tourists in full quarantine during COVID
'This is a photo of the Notre Dame, abandoned of its tourists in full lockdown during COVID.'[28]

28 Based on https://twitter.com/chouettephoto/status/1344600099113074691, retrieved March 18, 2022.

Though acceptability scores for *de* varied widely between speakers in our survey, this preposition seems to be slightly worse for most speakers in the (a) examples than in the (b) examples. It may be that better contexts can be constructed to make *de* more acceptable in the atelic (b) sentences. The difference in acceptability for *par* is very small, perhaps because *par* acts as the default Agent preposition.

1.3.5 French *de* and *par* in passives: discussion

In the previous subsections we have shown that the choice between *de* and *par* depends on several properties of the relation between the Agent and the event. The most important factors are bringing about a change (sections 1.3.1 and 1.3.2), volitionality (section 1.3.3), and telicity (section 1.3.4). When both *de* and *par* are possible, we found that the choice is influenced by one or more of these properties. The stative/dynamic contrast is a property of clauses that plays in the background of many sentences, but we are not aware of any sentences where the choice between *de* and *par* is only or primarily conditioned by this property.

It seems to us that the property of bringing about a change is primary: if a predicate can imply a change, the use of *par* will force it to do so. Thus, if a predicate *can* imply all of change, volitionality, and telicity, the use of *par* will imply change but not necessarily volitionality or telicity. As evidence for this, note that entrainment causation can be seen as change without telicity (Copley & Harley 2022), and that cases of entrainment causation with clear changes require *par* (recall [1.18] above). Similarly, non-volitional Agents that bring about a change still require *par* as well:

(1.34) *Le chien est lavé* ***par***/***de*** *Marie, bien qu'elle n'en*
the dog is washed *par*/ *de* Marie although=she NEG=of.them
avait pas envie[29]
had NEG desire
'The dog was washed by Marie, though she didn't want to (wash it).'

We conclude the following for sentences in which both *de* and *par* are allowed:

(1.35) a. If change could be implied by the event, *par* will imply change and *de* will imply lack of change. The scale to measure change may be

29 This example, built on (1.1a), was not included in our survey, but is uncontroversial.

inferred contextually. *Par* does not necessarily imply any other properties of proto-agentivity.

b. If change is excluded by the event, *par* will imply a higher level of volitionality and/or telicity than *de*.

The differences in presupposed proto-agentivity can be expressed very transparently with the approach to *by*-phrases proposed in section 1.2.2. We argued there that Agent prepositions have denotations of type $\langle e, \langle s, t\rangle\rangle$, and thus naturally lend themselves to express properties of the relation between the Agent and the event. Assuming the polymorphic denotations of *de* and *par* from (1.15a) and (1.16a), the concrete meanings in passives can be given as in (1.36). As discussed for (1.6b), we use Initiator(x, e) quite broadly here; our interest is in the difference in presupposed proto-agentivity.

(1.15a) $[\![\text{par}]\!]_{\langle \eta, \langle \theta, t\rangle\rangle} = \lambda g_\eta \lambda f_\theta . f$ is through/via g

(1.16a) $[\![\text{de}]\!]_{\langle \eta, \langle \theta, t\rangle\rangle} = \lambda g_\eta \lambda f_\theta . f$ is from/of g

(1.36) a. $[\![\text{par}_{\text{agentive}}]\!]_{\langle e, \langle s, t\rangle\rangle}$
$= \lambda x \lambda e . e$ is through/via x

interpretation: Initiator(x, e)

presupposed: x has high proto-agentivity in e

b. $[\![\text{de}_{\text{agentive}}]\!]_{\langle e, \langle s, t\rangle\rangle}$
$= \lambda x \lambda e . e$ is from/of x

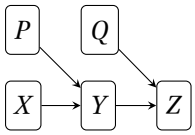
interpretation: Initiator(x, e)

presupposed: x has low proto-agentivity in e

Why would *de* 'from, of' imply low proto-agentivity and *par* 'through, via' high proto-agentivity? Ultimately, the answer depends on how humans conceptualize causation using spatial notions. We can only sketch the outline of a possible answer here. Consider again the notion of the causal chain in which arguments are either antecedent or subsequent to the Patient (section 1.2.1). Croft (2012: 222–226) showed that arguments antecedent to the Patient (Agent, Instrument, etc.) are typically marked by ablative or perlative prepositions ('from', 'through'), whereas arguments subsequent to the Patient (Beneficiary, Goal, etc.) are typically marked by allative prepositions ('to', 'for'). We suggest that the causal meaning of a preposition is not only determined by the relative position expressed by its spatial meaning, but also by the *distance* it expresses. In particular, *par* 'through, via' places the Figure at a smaller distance from the Ground than *de* 'from, of', and would

therefore be used for Agents at a smaller "causal distance" from the Patient. This smaller causal distance would then be interpreted as a greater ability for the Agent to affect the Patient, and hence, as a higher degree of proto-agentivity.[30] This argument based on the causal chain can be extended to other causal prepositions; for instance, *avec* 'with' expresses an even smaller distance than *par* and can be used for Instruments, which stand between the Agent and the Patient in the causal chain.

Another way to understand the difference in meaning between *par* and *de* uses causal models (Halpern & Pearl 2005). Causal models are directed graphs representing the dependency of variables on each other, as exemplified here:

(1.37) 

In this model, Z depends on X only through Y. It can be proven, but is intuitively clear, that the set of cases (i.e., variable assignments) in which Z depends on X is a subset of the set of cases in which it depends on Y. For example, the formula for Y may disregard the value of X for certain values of P; in this case, Z still depends on Y, but not on X.[31] Therefore, a greater distance between two variables in the causal model corresponds to a smaller dependency of the effect on the cause. If Z were to represent a scale on which change is measured, and X and Y represent actions by Agents or other causing arguments, a greater distance therefore corresponds to a smaller degree to which an Agent can affect the Patient. This is another way in which the link between the distance expressed by *de* 'from, of' can be related to the implication of low proto-agentivity in its causal uses.

30 It is important to note that in different languages, prepositions with very similar spatial meanings may have different causal meanings. As we will see in chapter 2 in more detail, in Biblical Hebrew, Causes marked by *min* 'from' are more "dominant" than those marked by *bə* 'in'. In the sense of dominance defined there, it seems that French *de* is intuitively *less* dominant than *par*, even though it has roughly the same spatial meaning as Biblical Hebrew *min*. I return to the possibility of multiple cognitive conceptualizations of the causal domain in chapter 6.

31 We leave more complex models out of consideration here (e.g., if Q were to also depend on X, so that Z depends on X through two paths). It is not clear that natural languages can describe such models without periphrasis.

1.4 Related work

In this section we discuss related work. Section 1.4.1 compares our results to previous work on French *de* and *par*, and section 1.4.2 discusses other formal accounts of *by*-phrases in passives. Finally, in section 1.4.3 we compare our approach to polysemy in causal prepositions using polymorphic types to an alternative using sum types.

1.4.1 Related work on French *de* and *par*

The distinction between *de* and *par* has received quite some attention in the literature. Though the choice depends in part on register (*de* being more formal; Gougenheim 1938: 307; and nowadays felt to be archaic), our focus is here on semantic distinctions. An intuitive approach based on the difference between verbal and adjectival passives cannot be used to describe the data.[32] Instead, the difference is usually framed in terms of Aktionsart, with *de* co-occurring with stative events (Zumthor & von Wartburg 1947: 297). The most complete descriptive generalization is given by Straub (1974):

(1.38) a. The Agent of a verb that denotes a non-state is always marked by *par*.

b. Verbs denoting states with animate Agents can be marked by both *de* and *par*.

c. Verbs denoting states with inanimate Agents always take *de*.

In our analysis we used the notion of proto-agentivity, claiming that *de* expresses low proto-agentivity, whereas *par* is used for more prototypical Agents, and also as a default. This derives the intuition of (1.38), if we remember that the stative/dynamic contrast is related to proto-agentivity through transitivity (Hopper & Thompson 1980), and that animacy is related to proto-agentivity through the notions of volitionality and bringing about a change.

32 Given the preference of *de* for statives, we might expect that adjectival passives take *de*, while verbal passives take *par*. However, some simple tests based on Hallman (2021) show that this idea does not pan out. *De*-passives can be verbal, too (cf. [1.2]), and combinations with adjectival morphology and coordination with adjectives do not rule out either *par*-phrases or *de*-phrases (*Le garçon est gentil et très adoré par le/du grand-père* 'The boy is kind and very much loved by the grandfather'), nor do verbs like *sembler* 'seem' (*Le garçon semble adoré par le/du grand-père* 'The boy seems loved by the grandfather'). It is not clear that the verbal-adjectival passive distinction is useful in French.

However, (1.38) is not precise enough, since many of the judgments from section 1.3 are incompatible with it.[33] As just one example, (1.25b) is typically seen as a state and has an inanimate Agent, yet allows *par* (contra [1.38c]):

(1.25b) *Le village est entouré* $^{0.79}$***d'*** /$^{0.77}$***par*** *une chaîne de montagnes,*
the village is surrounded *de/* *par* a chain of mountains
à cause de laquelle les services d'urgence arrivent toujours
at cause of which the services of=emergency arrive always
trop tard
too late
'The village is surrounded by a mountain chain, because of which the emergency services always arrive too late.'

We explained this by appealing to the notion of entrainment causation (section 1.3.1) to make a more precise distinction than that between "states" and "non-states", and by allowing for contextually inferred scales to measure change (section 1.3.2).

Another problem with (1.38) is that it does not predict anything regarding the choice between *de* and *par* when both are possible (1.38b). We resolved this by moving away from a strict rule-based approach ("if a sentence has these properties, this preposition must be used") to a more flexible approach based on the *degree* of proto-agentivity. This approach also does more justice to the variation between speakers and to the fact that for many sentences the difference in acceptability between the two prepositions is small.

1.4.2 Related work on *by*-phrases

There are two mainstream formal semantic accounts of *by*-phrases in passives. The main difference between them is whether the *by*-phrase is an argument or an adjunct.[34] The approach on which we built our own proposal in section 1.2.2 is that of Angelopoulos et al. (2020), who argue that

33 Many examples are also discussed by Gaatone (1998: 175–210). Our analysis is compatible with his data, but we do not systematically compare our work to his since he does not propose an explanatory theory.

34 See Williams (2015: 281–291). We focus on what he terms "Base Argument Theories", which assume that "some syntactic part of a short passive clause, and some part of the host in a long passive, has a functional semantic argument in the deep-S role" (Williams

the *by*-phrase is an argument of v (the head of the light verb phrase introducing the Agent). It thus takes the same place as the external argument in an active sentence. Angelopoulos et al. (2020) are not explicit about a formal semantic analysis but base themselves on Collins (2018), who gives the Agent preposition the identity function as its denotation:

(1.39) $[\![\text{by}_{\text{agentive}}]\!] = \lambda x.x$ (Collins 2018: 4)

As a result, the denotation of the *by*-phrase is of type $\langle e \rangle$ and can therefore compose by Function Application with the denotation of v', which has type $\langle e, \langle s, t \rangle \rangle$. By contrast, we argued that the type of *by* must be $\langle e, \langle s, t \rangle \rangle$ (and composes with v' using Event Identification). Section 1.2.2 presented this mostly as following from our suggestion for the formalization of principled polysemy in section 1.2.1, but there is an independent reason why we believe (1.39) is not ideal. With the denotation in (1.39), *by* is essentially seen as a kind of case marker, needed to mark the argument but semantically vacuous. This may suffice for English *by*, but we have seen in section 1.3 that French *de* and *par* are *not* semantically vacuous. Instead, they carry a presupposition concerning properties of the Agent's relation to the event.

Since this presupposition does not concern an inherent property of the Agent (e.g., animacy) but a property of the relation of the Agent to the event, we believe the most transparent way to do this is to have the denotation of *by* take the event as an argument. It would be much less transparent to have a presupposition "x has high proto-agentivity in e" on a preposition with the meaning in (1.39), since it would have to be contextually inferred what event e refers to. We find a formalization in which the presupposition only depends on variables provided as arguments to the denotation preferable. Admittedly, this argument does not entirely rule out an analysis along the lines of (1.39). However, we see no immediate benefit to such an analysis, while we do believe it is worthwhile to make the reference of the presupposition to the event transparent and to systematically derive the meaning of Agent prepositions from a more general, polymorphic meaning of that preposition (as discussed in section 1.2.1).

Another common approach to *by*-phrases is found in Bruening (2013) (Legate 2014 has a similar proposal). Bruening (2013) assumes a VoiceP of type $\langle e, \langle s, t \rangle \rangle$. In a regular active sentence, the $\langle e \rangle$ argument of this projec-

2015: 282). We do not discuss No Base Argument Theories, being unaware of formal semantic analyses in such theories.

tion is saturated by the external argument (1.40a), and in a passive without a *by*-phrase, it is saturated by existential closure of an additional Pass projection above VoiceP (1.40b). In a passive with a *by*-phrase, the *by*-phrase is an adjunct to Voice′. It is seen as a purely functional element of type $\langle e, \langle \langle e, \langle s, t \rangle \rangle, \langle s, t \rangle \rangle \rangle$ (1.40c), which fills in the argument of Voice (i.e., it performs the same task as the external argument in the active voice); the passive Voice head is semantically vacuous (1.40d):

(1.40) a. $[\![\text{Voice}]\!]_{\langle \langle s,t \rangle, \langle e, \langle s,t \rangle \rangle \rangle} = \lambda p \lambda x \lambda e.p(e)$ & Initiator(e, x)
(Bruening 2013: 21)

b. $[\![\text{Pass}]\!]_{\langle \langle e, \langle s,t \rangle \rangle, \langle s,t \rangle \rangle} = \lambda p \lambda e. \exists x\text{:}\, p(x, e)$
(without *by*-phrase; Bruening 2013: 25)

c. $[\![\text{by}]\!]_{\langle e, \langle \langle e, \langle s,t \rangle \rangle, \langle s,t \rangle \rangle \rangle} = \lambda x \lambda p \lambda e.p(x, e)$ (Bruening 2013: 25[35])

d. $[\![\text{Pass}]\!]_{\langle \langle s,t \rangle, \langle s,t \rangle \rangle} = \lambda p \lambda e.p(e)$ (with a *by*-phrase; Bruening 2013: 25)

In this analysis, the denotation of the Agent preposition has access to the event argument, so our critique of Collins (2018) does not apply. However, note that this approach makes some unusual syntactic assumptions (cf. Roberts 2019: 437), and also fails to account for certain binding facts (Collins 2018; Angelopoulos et al. 2020): the Agent in a *by*-phrase can bind an anaphor in the VP (1.41a), which is expected if the *by*-phrase is an argument, as in Angelopoulos et al. (2020) and our modification of it, but unexpected if the *by*-phrase is an adjunct. Both *de* and *par* behave like English *by* with respect to binding (1.41b), in contrast to other French prepositions (1.41c):

(1.41) a. *The packages were sent* ***by*** *the children$_i$ to themselves$_i$.*
(Angelopoulos et al. 2020: 11)

b. *Les enfants vont jouer au foot accompagnés* ***de***/***par*** *leurs*
the children go play_soccer accompanied *de*/*par* their
parents$_i$ conformément à leur$_i$ propre volonté
parents according to their own will

'The children are going to play soccer accompanied by their parents according to their own wish.'

35 This denotation differs slightly from Bruening's ($[\![\text{by}]\!] = \lambda x \lambda p \lambda e.p(e, x)$, with the arguments to p swapped), for consistency with the rest of the chapter.

c. *Les enfants vont jouer au foot* ****avec/*sans /*chez/*pour*** *leurs*
the children go play_soccer with/ without/ at / for their
parents$_i$ conformément à leur$_i$ propre volonté
parents(') according to their own will
'The children are going to play soccer with/without/at/for their parents(') according to their own wish.'

For this reason, we adopted an account under which the *by*-phrase is an argument in section 1.2.2.[36]

1.4.3 Polymorphism compared to sum types

Finally, in this subsection we discuss an alternative approach to the polysemy of causal prepositions, which uses the notion of sum types from Type Composition Logic (Asher 2011). Maienborn & Herdtfelder (2017) show that the German preposition *von* 'from, of, by' can be used for causal adjuncts with both stative and eventive readings, which have different inferential properties. Stative (1.42a) implies that the hailstones are (i) on the square and (ii) white, while eventive (1.42b) does not imply that the shoes are (i) still on the bench or (ii) dirty.

(1.42) a. *Der Platz ist weiß* ***von*** *den Hagelkörnern*
the square is white *von* the hailstones
'The square is white from the hailstones.'
(Maienborn & Herdtfelder 2017: 285)

b. *Die Bank ist dreckig* ***von*** *den Schuhen*
the bench is dirty *von* the shoes
'The bench is dirty from the shoes.'
(Maienborn & Herdtfelder 2017: 285)

Maienborn & Herdtfelder (2017) propose an approach in which these inferential properties are derived from a different type. They suggest that stative *von* expresses a causal relation between TROPEs (Moltmann 2007), while eventive *von* expresses a relation between EV(ent)s. Simplifying some matters that are inconsequential to our comparison here, arguments of *von* are of the sum type EV ⊔ TROPE: they are either of type EV, or of type TROPE.

36 Ingason et al. (2016) present a variation on Bruening (2013) using the identity function as the denotation of the Agent preposition. It thus suffers from both the problem with the identity function approach *and* the problem with the binding facts.

Since the actual type of the argument propagates, inferential differences can be derived from whether the argument is an event or a trope.

This approach is superficially similar to ours with a polymorphic type $\langle \eta, \langle \theta, t \rangle \rangle$ (1.5): one might say that EV and TROPE are two types with which η and θ can be instantiated. However, note that Maienborn & Herdtfelder (2017) are only concerned with uses of *von* in causal adjuncts. The type constraint does not generalize to Agentive *von*, let alone meanings in other domains such as that of space (*von hinten* 'from behind') or time (*von morgens* 'from morning'). The type constraint could be modified to include more possible argument types, but this would lead to a rather complex lexical entry. In order to describe the full range of uses of a preposition, it seems preferable to us to separate its domain-specific meaning from its general meaning, as in the approach developed in (1.5). This enables the polymorphic typing approach developed here to capture this polysemy efficiently. The requirement spelled out in section 1.2 that each concrete interpretation is motivated by a cognitive linguistic explanation prevents the model from over-generating.

Besides being more minimal, an approach in which a preposition's general (polymorphic) meaning is separate from domain-specific (concrete) interpretations is also in line with other observations concerning the polysemy of prepositions. We have already seen that different languages may have different causal interpretations of what seems to be the same spatial preposition. In an approach where general meaning is separate from concrete interpretations, we can hypothesize that in such cases the prepositions have the same general meaning, even though their concrete interpretations may differ from language to language. The overlap in concrete spatial meaning is then easily explained, while the differences in causal meaning are due to different conceptualizations of causation. In other approaches (either with a single complex lexical entry, or multiple, unrelated entries) the overlap in meaning between such prepositions would be coincidental.

1.5 Conclusions

Common approaches to *by*-phrases in passives treat the Agent preposition as semantically vacuous: it only rearranges the arguments so that the argument of *by* fulfills the same role as the external argument in the correspond-

ing active sentence (Bruening 2013; Legate 2014; Collins 2018; Angelopoulos et al. 2020). This chapter has put forward three arguments against this view.

First, cross-linguistic research shows that Agent prepositions develop from prepositions with specific spatial meanings, and cognitive linguistic arguments can be given to relate these spatial meanings to the function of Agent marking (Croft 2012: 222–226). However, common approaches to *by*-phrases essentially treat agentive *by* as accidentally homonymous with spatial *by*, and therefore cannot explain this cognitively motivated cross-linguistic pattern.

Second, we discussed languages with more than one Agent preposition, where the choice of the Agent preposition is semantically motivated. Building on Straub (1974) and others, we showed that French *de* 'from, of, by' is used for Agents with low proto-agentivity, whereas *par* 'through, by' is the default Agent preposition and used for Agents with high proto-agentivity. Current approaches to *by*-phrases may be able to express such differences but are, we feel, not the most transparent way to do so.

Third, common approaches to *by*-phrases in passives do not generalize to other syntactic environments in which the same preposition appears with a causal meaning. This is especially problematic in the case of French, where it can be shown that the meanings of *de* and *par* in passives are similar to those in causal adjuncts. In causal adjuncts, *de* is related to stativity (marking causes that are situations), while *par* is related to dynamicity (marking causes that are forces). Stativity and dynamicity are related to low and high proto-agentivity, respectively, which is what *de* and *par* presuppose in passives. This parallel suggests that a formalization must not be limited to the syntactic environment of the passive.

The alternative we present builds on the notion of principled polysemy (Tyler & Evans 2003): the idea that the many different meanings of prepositions are not accidentally homonymous, but are instead related to each other through a shared core. We propose to formalize this using a polymorphically typed general denotation. This general denotation is typically an abstract form of the spatial meaning of the preposition, since spatial meanings tend to be original in processes of semantic extension. It has a polymorphic type: $\langle \eta, \langle \theta, t \rangle \rangle$, in which η and θ still have to be instantiated with concrete types to obtain a concrete interpretation.

In this system, the exact meaning of a preposition in context will depend on three things. First, the syntactic and semantic context forces a certain

type on the denotation of the preposition. Second, the interpretation is restricted to a certain domain depending on the type (e.g., the spatial domain for $\langle e, \langle e, t\rangle\rangle$ or the causal domain for $\langle e, \langle s, t\rangle\rangle$). Third, the concrete meaning within that domain depends on the way that domain is spatially conceptualized in the mind. For example, in the case of causation, causes are typically conceived of as antecedent and/or proximate to effects (e.g., Croft 2012: 222–226), which can explain why prepositions like English *by* and French *de* and *par* receive the causal interpretation they do.

This approach can be extended to systems with more types than e, s, and t to derive the meanings of different prepositions in a broader range of contexts. In section 1.2.3 we showed how this might work, analyzing French *de* and *par* in causal adjuncts in the force-theoretic framework of Copley & Harley (2022). We hope that the approach to the polysemy of prepositions put forward here is useful for other prepositions in other domains as well.

1.6 Survey data

Most of the French example sentences from this chapter are based on real-world examples on the web or examples from the literature. Examples from the web were found through Google and Linguee and come from sources that we assumed were written by native speakers. We adapted sentences to add context to promote a certain reading and added sentences to create minimal pairs. Examples (1.2) and (1.30) were unintentionally ambiguous.

We confirmed our judgments, except for some uncontroversial examples, with a small number of native speakers in an informal survey. We invited informants whom we expected to still be familiar with a more formal or archaic register, based on age, education level, and religious background (as Bible translations tend to use a more conservative register). 21 Informants from France, Belgium, and Switzerland completed the survey, with a mean age of 49 (standard deviation 19); 16 (76%) had at least a Master's degree. Though the sample size is not large enough to expect statistically significant results, the tendencies in the data align with our own judgments. For each sentence there was also space for comments, for example to remark on differences in interpretation when respondents considered both *de* and *par* were acceptable. These comments were all in line with our own intuitions.

Each sentence was presented as-is to the participants without any additional context. We replaced the Agent preposition with a blank (e.g., *Le chien est lavé … Marie* for [1.1a]). Participants were then asked to rate the acceptability of both *de* and *par* on a 6-point Likert scale ranging from *pas du tout acceptable* 'not at all acceptable' to *parfaitement acceptable* 'perfectly acceptable'. They were asked to rate a preposition as acceptable if they were familiar with its use in the given context even if they would not use it themselves.

The results are summarized in tables 1.1 and 1.2 below, in the order the sentences are discussed in sections 1.2 and 1.3. The table indicates for each sentence whether we expected the sentence to have relatively high or proto-agentivity (and thus, whether we expected *par* or *de*, respectively, to be more acceptable). The scores for each preposition are presented in stacked bar charts. They were also recoded to values from −1 (not at all acceptable) to 1 (perfectly acceptable) to be able to compute the mean, which we only used to give a quick impression of the general tendency in judgment marks throughout this chapter.

Example			Distribution	Mean
Causal adjuncts (section 1.2.3)				
(1.14a)	*mort*[1]	*de*		1.00
		par		−0.87
(1.14b)	*cassée*[2]	*de*		−0.90
		par		0.37
Change (section 1.3.1)				
(1.1a)	*lavé* (high)	*de*		−0.96
		par		1.00
(1.18)	*maintenu* (high)	*de*		−0.81
		par		1.00
(1.20a)	*suivie* (high)	*de*		−0.45
		par		1.00
(1.20b)	*traversé* (high)	*de*		−0.77
		par		0.96
Contextually implied change (section 1.3.2)				
(1.21a)	*adorée* (low)	*de*		0.60
		par		0.77
(1.21b)	*adorée* (high)	*de*		0.50
		par		0.92
(1.22a)	*aimé* (low)	*de*		0.92
		par		0.70
(1.22b)	*aimé* (high)	*de*		0.83
		par		0.77
(1.23a)	*respecté* (low)	*de*		0.77
		par		0.89
(1.23b)	*respecté* (high)	*de*		0.64
		par		0.81
(1.2)	*accompagnés*	*de*		0.87
		par		0.94
(1.24a)	*accompagné* (high)	*de*		0.79
		par		0.96
(1.24b)	*accompagné* (low)	*de*		0.79
		par		0.89
(1.25a)	*entouré* (low)	*de*		0.87
		par		0.73
(1.25b)	*entouré* (high)	*de*		0.79
		par		0.77

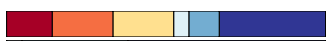
(*continued on the following page*)

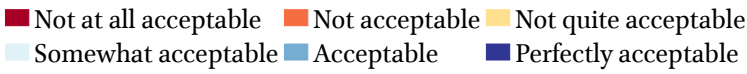
(*continued from the previous page*)

Volitionality (section 1.3.3)

(1.26a)	*précédé* (low)	*de*		0.83
		par		0.50
(1.26b)	*suivi* (high)	*de*		0.35
		par		0.98
(1.27a)	*suivi* (low)	*de*		0.98
		par		0.31
(1.27b)	*suivi* (high)	*de*		0.77
		par		0.64
(1.28a)	*suivis* (low)	*de*		0.89
		par		0.50
(1.28b)	*suivis* (high)	*de*		0.77
		par		0.68
(1.29a)	*suivie* (low)	*de*		0.81
		par		0.71
(1.29b)	*suivie* (high)	*de*		0.73
		par		0.77
(1.30)	*suivi* (high)	*de*		0.66
		par		0.98

Telicity (section 1.3.4)

(1.32a)	*abandonnés* (high)	*de*		−0.12
		par		1.00
(1.32b)	*abandonnés* (low)	*de*		−0.07
		par		0.92
(1.33a)	*délaissée* (high)	*de*		0.10
		par		0.83
(1.33b)	*délaissé* (low)	*de*		0.28
		par		0.85



1 Tested with *faim*.
2 Tested with *l'impact du ballon*.

Table 1.1 Survey data (high = high proto-agentivity; low = low proto-agentivity).

Example		*de*						*par*					
		−−−		...		+++		−−−		...		+++	
(1.14a)	*mort*	0	0	0	0	0	21	15	5	1	0	0	0
(1.14b)	*cassée*	16	5	0	0	0	0	4	2	1	0	2	12
(1.1a)	*lavé* (high)	19	2	0	0	0	0	0	0	0	0	0	21
(1.18)	*maintenu* (high)	13	7	0	1	0	0	0	0	0	0	0	21
(1.20a)	*suivie* (high)	8	3	6	2	2	0	0	0	0	0	0	21
(1.20b)	*traversé* (high)	13	5	2	1	0	0	0	0	0	0	2	19
(1.21a)	*adorée* (low)	0	0	1	4	10	6	0	1	0	2	4	14
(1.21b)	*adorée* (high)	0	1	3	3	7	7	0	0	0	1	2	18
(1.22a)	*aimé* (low)	0	0	0	1	2	18	0	0	2	3	4	12
(1.22b)	*aimé* (high)	0	0	2	0	3	16	0	0	1	1	7	12
(1.23a)	*respecté* (low)	0	0	1	0	9	11	0	0	0	0	6	15
(1.23b)	*respecté* (high)	0	1	2	2	5	11	0	1	0	1	4	15
(1.2)	*accompagnés*	1	0	0	0	2	18	0	0	0	0	3	18
(1.24a)	*accompagné* (high)	0	1	1	1	2	16	0	0	0	0	2	19
(1.24b)	*accompagné* (low)	0	1	1	2	0	17	0	1	0	0	2	18
(1.25a)	*entouré* (low)	1	0	0	0	2	18	1	0	1	1	4	14
(1.25b)	*entouré* (high)	0	0	1	2	4	14	0	1	0	2	4	14
(1.26a)	*précédé* (low)	0	1	0	1	3	16	1	0	1	7	4	8
(1.26b)	*suivi* (high)	1	3	1	5	4	7	0	0	0	0	1	20
(1.27a)	*suivi* (low)	0	0	0	0	1	20	1	3	3	3	4	7
(1.27b)	*suivi* (high)	0	1	1	1	3	15	0	1	2	3	3	12
(1.28a)	*suivis* (low)	0	0	1	0	3	17	1	1	1	5	4	9
(1.28b)	*suivis* (high)	0	0	0	3	6	12	0	0	3	2	4	12
(1.29a)	*suivie* (low)	0	0	1	1	5	14	1	0	2	0	4	14
(1.29b)	*suivie* (high)	0	1	1	3	1	15	0	0	0	4	4	13
(1.30)	*suivi* (high)	0	1	1	4	3	12	0	0	0	0	1	20
(1.32a)	*abandonnés* (high)	3	7	3	3	1	4	0	0	0	0	0	21
(1.32b)	*abandonnés* (low)	3	5	3	5	2	3	0	0	1	0	1	19
(1.33a)	*délaissée* (high)	3	4	4	1	2	7	0	0	0	1	7	13
(1.33b)	*délaissé* (low)	0	5	2	4	4	6	0	0	1	1	3	16

Table 1.2 Raw data for table 1.1 (−−− = not at all acceptable; +++ = perfectly acceptable).

The distinction between causal *min* 'from' and *bə* 'in' 2

Abstract In Biblical Hebrew, both *min* 'from' and *bə* 'in' are used to mark causing arguments (Agents, Instruments, Reasons, …). Reference works list the thematic roles each preposition can mark, but do not address the differences between the two. We argue that the contrast is one of "dominance": *min*-causers are more dominant than *bə*-causers. They can fully determine the effect, whereas the effect of a *bə*-causer can be altered or prevented. This distinction derives from the spatial meanings of these prepositions based on an abstract spatial representation of the causal domain. The object of *min* is a Source or Origin, which is interpreted as being the instigator of a causal chain, and thus having dominance over that chain from instigation to effect. By contrast, the Locative preposition *bə* describes a location close to a Ground, which is interpreted as being able to cause an effect, but not necessarily in a dominant way.

2.1 Introduction

When two prepositions can be used in the same context, the choice between them often depends on subtle differences in meaning. Reference works tend to have fairly in-depth descriptions of Biblical Hebrew prepositions

This chapter was originally published as: Camil Staps & Martijn Beukenhorst. 2024. Semantic properties of prepositions: The distinction between causal *min* 'from' and *b*^{e} 'in'. *Brill's Journal of Afroasiatic Languages and Linguistics* 16(2). DOI: 10.1163/18776930-01602002. Camil Staps and Martijn Beukenhorst designed the study. Camil Staps collected the data, performed the analysis, and wrote the article, in regular discussion with Martijn Beukenhorst.

We thank Ellen van Wolde, Johan Rooryck, Benjamin Suchard, and two anonymous reviewers for comments on earlier versions of this article. We are also grateful to the audience of the Linguistics and Biblical Hebrew/NAPH session at the SBL Annual Meeting 2022 in Denver, Colorado, as well as to the audience of the COCOA seminar, for their useful feedback. Any mistakes are ours alone.

individually, but the contrasts between them are rarely made explicit. As a result, it often remains hard to articulate why the choice for a particular preposition in a specific text was made. This chapter aims to address part of this gap by looking at the prepositions *bə* 'in' and *min* 'from'. In particular, we are concerned with the causal uses of these prepositions, as exemplified in the following examples:[1]

(2.1) Josh. 10:11: רַבִּים אֲשֶׁר־מֵתוּ בְּאַבְנֵי הַבָּרָד מֵאֲשֶׁר הָרְגוּ בְּנֵי יִשְׂרָאֵל בֶּחָרֶב׃

rabb-īm	*ʾăšɛr*	*mēṯ-ū*	***bə***=*ʾaḇn-ē*	*hab=bå̄rå̄ḏ*	*mē=ʾăšɛr*
many-PL	REL	die\PFV-3PL	**in**=stone-PL.of	the=hail	from=REL

hå̄rəḡ-ū	*bən-ē*	*yiśrå̄ʾēl*	***bɛ***=*ḥå̄rɛḇ*
kill\PFV-3PL	son-PL.of	Israel	**in**=sword

'There were more who died *because of* the hailstones than the sons of Israel killed *with* the sword.' (ESV)

(2.2) 1 Sam. 28:20: וַיְמַהֵר שָׁאוּל וַיִּפֹּל מְלֹא־קוֹמָתוֹ אַרְצָה וַיִּרָא מְאֹד מִדִּבְרֵי שְׁמוּאֵל

wa-y-əmahēr-Ø	*šå̄ʾūl*	*way-y-ippōl-Ø*	*məlōʾ*	*qōmå̄ṯ-ō*	*ʾarṣ-å̄*
WAYQ-3M-hurry-SG	Saul	WAYQ-3M-fall-SG	filled.of	height-his	ground-ALL

way-y-irå̄ʾ-Ø	*məʾōḏ*	***mid***=*diḇr-ē*	*šəmūʾēl*
WAYQ-3M-fear-SG	very	**from**=word-PL.of	Samuel

'Then Saul fell at once full length on the ground, filled with fear *because of* the words of Samuel.' (ESV)

In (2.1), *bə* is used twice to mark the Instrument used by another entity.[2] In (2.2), *min* is used to mark the Reason for Saul's fear. But *min* can also be used to mark Instruments, as seen in (2.3), and *bə* can be used to mark Reasons (2.4):

(2.3) 2 Sam. 7:29: וְעַתָּה הוֹאִל וּבָרֵךְ אֶת־בֵּית עַבְדְּךָ ... וּמִבִּרְכָתְךָ יְבֹרַךְ בֵּית־עַבְדְּךָ לְעוֹלָם׃

1 I provide translations from the ESV as a reasonably literal version that is also accessible to a modern audience unfamiliar with the source texts.

2 The capitalized terms Agent, Cause, Instrument, and Reason represent thematic roles (Davis 2011; Harley 2011; Primus 2016). We only use these roles to give a quick impression of the functions of these prepositions, and replace thematic roles with more precisely defined notions in section 2.3. When discussing secondary literature, we only capitalize these terms when it is clear that the author sees them as thematic roles.

wə=ʿattå̄ *hōʾēl-Ø* *ū=bå̄rēḵ-Ø* *ʾɛṯ* *bēṯ*
and=now agree\IMP-M.SG and=bless\IMP-M.SG OBJ house.of
ʿaḇdə-ḵå̄ ... *ū=**mib**=birḵå̄ṯ-əḵå̄* *y-əḇōraḵ-Ø*
servant-yours ... and=**from**=blessing-yours 3M-bless\PASS.IPFV-M.SG
bēṯ *ʿaḇdə-ḵå̄* *lə=ʿōlå̄m*
house(M).of servant-yours to=eternity

'Now therefore may it please you to bless the house of your servant, ..., and *with* your blessing the house of your servant shall be blessed forever.' (ESV)

(2.4) Gen. 41:36: וְלֹא־תִכָּרֵת הָאָרֶץ בָּרָעָב׃

wə=lō' *ṯ-ikkå̄rēṯ-Ø* *hå̄=ʾå̄rɛṣ* ***b**=å̄=rå̄ʿå̄ḇ*
and=not F-cut_off\MID.IPFV-3SG the=land(F) **in**=the=famine

'(That food shall be a reserve ...), so that the land may not perish *through* the famine.' (ESV)

The reference works describe the arguments of these prepositions with terms like Instrument, Cause, and Agent, but often lack a clear working definition of such thematic roles.[3] Furthermore, the examples above show that there is overlap between the functions of *bə* and *min*. The literature is by and large silent on the precise factors conditioning the choice between them.

This chapter argues that there is a difference in the degree of "dominance" that the argument has over the situation: *min*-causers are fully dominant, or nearly so, while *bə*-causers are less dominant. In the examples above, this works as follows. In (2.1), the hailstones and the sword are manipulated by other entities (God and the Israelites, respectively) to bring about the event. These other entities are in control of the situation; *bə* is used because the hailstones and the sword participate in the event but could not, as inanimate entities, prevent it from happening or otherwise change the outcome. In (2.2), with *min*, Saul is completely overcome by fear of the words of Samuel; those words have taken full control over him and are the only reason for his current state of mind. In (2.3), *min* is used to mark the blessing as dominant over other, hypothetical, intervening causers: it is even so powerful that it lasts forever, no matter what other events may occur. Finally, in (2.4), *bə* is used to downplay the famine as only a minor influence; after all, Egypt has prepared for this famine by storing up food: the famine is here

3 The exact description varies; see section 2.2 for details. The main reference works consulted are Gesenius (1910), Waltke & O'Connor (1990), Joüon & Muraoka (2006), Van der Merwe et al. (2017), and, specifically on *bə*, Jenni (1992).

dominated by an intervening cause (the storing of food by the Egyptians); as a result, the land does not perish.

We will show below how this notion of "dominance" can be formalized using causal models (Pearl 2000; Halpern & Pearl 2005, among others). We also argue that this semantics can be derived from the spatial meaning of the prepositions, explaining how the spatial notion of distance is interpreted in the causal domain. We first summarize how *min* and *bə* are usually described (section 2.2). Section 2.3 discusses the relevant theoretical background on causation and proposes a formal definition of the notion of "dominance". In section 2.4 we show how this definition accounts for the distribution of *min* and *bə* in environments where both are possible. Section 2.5 briefly discusses the status of the difference in meaning between *min* and *bə* by investigating the behavior under negation, and section 2.6 concludes.

2.2 Current descriptions of *min* and *bə*

There is quite some discussion in the reference works as to the different causal functions of *min* and *bə*.[4] Waltke & O'Connor (1990: §11.2.5) distinguish three causal uses of *bə*. The division depends on inherent properties of the argument: inanimates are instruments (2.5–2.6), animates are agents (2.7), and reasons or originating forces are causes (2.8):

(2.5) Mic. 4:14: בַּשֵּׁבֶט יַכּוּ עַל־הַלְּחִי אֶת שֹׁפֵט יִשְׂרָאֵל׃

b=*aš*=*šēḇɛṭ*	*y-akk-ū*	*ʿal*	*hal*=*ləḥī*	*ʾēṯ*	*šōpēṭ-Ø*	
in=the=rod	3M-strike\IPFV-PL	on	the=cheek	OBJ	judge\PTCP-M.SG.of	*yiśrɑ̊ʾēl* Israel

'*with* a rod they strike the judge of Israel on the cheek' (ESV)

4 Many works also discuss a causal meaning of *lə* 'to', but the examples are not convincing and will not be discussed here. These cases are problematic since they can be read as 'in relation to', a more common meaning of *lə*, as in Gen. 31:15: 'Aren't we considered foreigners *by/in relation to* him?' It seems that agency or causation is an epiphenomenon at best, and not contributed by the preposition proper. See Gesenius (1910: §119f); Waltke & O'Connor (1990: §11.2.10g); Joüon & Muraoka (2006: §132f, 133d); Van der Merwe et al. (2017: §39.11.6df); Jenni (2000: 299–300).

(2.6) 1 Kgs. 1:40: וַתִּבָּקַע הָאָרֶץ בְּקוֹלָם׃

wat-t-ibbå̄qaʿ-Ø *hå̄=ʾå̄reṣ* ***bə***=*qōl-å̄m*
WAYQ-F-split\MID-3SG the=earth(F) **in**=noise-theirs

'so that the earth was split *by* their noise' (ESV)

(2.7) Gen. 9:6: שֹׁפֵךְ דַּם הָאָדָם בָּאָדָם דָּמוֹ יִשָּׁפֵךְ

šōp̄ēk̲-Ø *dam* *hå̄=ʾå̄ḏå̄m* ***b***=*å̄=ʾå̄ḏå̄m* *då̄m-ō*
pour_out\PTCP-M.SG.of blood.of the=man **in**=the=man blood(M)-his

y-iššå̄p̄ēk̲-Ø
3M-pour_out\MID.IPFV-SG

'Whoever sheds the blood of man, *by* man shall his blood be shed' (ESV)

(2.8) Gen. 18:28: הֲתַשְׁחִית בַּחֲמִשָּׁה אֶת־כָּל־הָעִיר

hă=ṯ-ašḥīṯ-Ø ***ba***=*ḥămiššå̄* *ʾɛṯ* *kål* *hå̄=ʿīr*
Q=2M-destroy\IPFV-SG **in**=five OBJ whole.of the=city

'Will you destroy the whole city *for* lack of five?' (ESV)

The causal categories of *min* are very similar: this preposition, too, can mark causes and means (2.9–2.11), as well as agents (2.12) and reasons for fear (2.13–2.14) (Waltke & O'Connor 1990: §11.2.11d).

(2.9) Ezek. 28:18: מֵרֹב עֲוֹנֶיךָ בְּעֶוֶל רְכֻלָּתְךָ חִלַּלְתָּ מִקְדָּשֶׁיךָ

mē=*rōḇ* *ʿăwōn-ɛ-k̲å̄* *bə*=*ʿɛwɛl* *rək̲ullå̄ṯ-ək̲å̄*
from=multitude.of iniquity-PL-yours in=injustice.of[5] trade-yours

ḥillal-tå̄ *miqdå̄š-ɛ-k̲å̄*
profane\PFV-2M.SG sanctuary-PL-yours

'*By* the multitude of your iniquities, in the unrighteousness of your trade, you profaned your sanctuaries' (ESV)

(2.10) 2 Sam. 3:37 (see section 2.4.3): כִּי לֹא הָיְתָה מֵהַמֶּלֶךְ לְהָמִית אֶת־אַבְנֵר בֶּן־נֵר׃

kī *lōʾ* *hå̄yə-ṯå̄* ***mē***=*ham=mɛlɛk̲* *lə=hå̄mīṯ* *ʾɛṯ* *ʾaḇnēr*
COMP not be\PFV-3F.SG **from**=the=king to=die\CAUS.INF OBJ Abner

bɛn *nēr*
son.of Ner

'... that it had not been the king's *will* (lit.: *from* the king) to put to death Abner the son of Ner' (ESV)

5 We take *bə* here as circumstantial: '*during* the unrighteousness ...' (*pace* Jenni 1992: 145).

(2.11) Gen. 9:11: וְלֹא־יִכָּרֵת כָּל־בָּשָׂר עוֹד מִמֵּי הַמַּבּוּל

wə=lō' *y-ikkārēṯ-Ø* *kāl* *bāśār* *ʿōḏ* ***mim***=*m-ē*
and=not 3M-cut_off\MID.IPFV-SG all.of flesh(M) again **from**=water-of
ham=mabbūl
the=flood

'… that never again shall all flesh be cut off *by* the waters of the flood' (ESV)

(2.12) Lev. 21:7: אִשָּׁה גְּרוּשָׁה מֵאִישָׁהּ

'iššā *gərūš-ā* ***mē***=*'īš-āh*
woman divorce\PTCP.PASS-F.SG **from**=man-hers

'a woman divorced *from* her husband' (ESV)

(2.13) Ps. 27:1: מִמִּי אִירָא

mim=*mī* *'-īrā'*
from=who 1SG-fear\IPFV

'whom shall I fear?' (ESV)

(2.14) Exod. 34:30: וְהִנֵּה קָרַן עוֹר פָּנָיו וַיִּירְאוּ מִגֶּשֶׁת אֵלָיו

wə=hinnɛ *qārān-Ø* *ʿōr* *pānn-āyw* *way-y-īrə'-ū*
and=behold shine\PFV-3M.SG skin(M).of face-his WAYQ-3M-fear-PL
mig=*gɛšɛṯ* *'ēl-āyw*
from=approach\INF to-him

'and behold, the skin of his face shone, and they were afraid *to* come near him' (ESV)

Waltke & O'Connor (1990) do not address the question what the difference between *bə* and *min* is, if both of them are used to mark instruments/means, agents, and reasons.[6] As mentioned before, this is true for all the reference works we consulted.[7] The discussion in the reference works seems

6 We do not find it plausible that the two prepositions are simply "interchangeable", apparently without any difference in meaning and distribution. This is what Haber (2009) seems to suggest (we thank Tania Notarius for this reference). For instance, Haber explains the use of *min* in Prov. 5:18 as a "valid late interchange". The fact that the meanings of the prepositions remain clearly distinct in most instances suggests that the occasional "interchanges" we do see cannot be random.

7 For Joüon & Muraoka (2006) the prepositions have roughly the same meaning as for Waltke & O'Connor (1990): *bə* marks instrument or means, instrumental cause, or plain cause with infinitives; they also mention that verbs of 'rejoicing in' can be seen as causal (Joüon & Muraoka 2006: §133c). *Min* expresses cause, source, or origin (Joüon & Muraoka 2006: §133e). In discussing the opposition between causal *bə* and *min*, they write

to be hampered by the use of thematic roles like Agent and Instrument, often without a rigorous definition. For example, Joüon & Muraoka (2006: §132d) suggest that the cause in (2.11) is "only instrumental" (presumably with God setting the event in motion). Perhaps the hesitance to read the waters in (2.11) as an ultimate causer rather than an instrument is based on the assumption that ultimate causers must be animate (cf. Bicknell 1984: 44). Modern theoretical linguistic work has suggested, however, that many phenomena that seem to be restricted by animacy are in fact restricted by teleological capability: the inherent ability of the entity to participate in the eventuality (Folli & Harley 2008). Natural forces, like the flood in (2.11), are textbook examples of entities that are conceived of as producing energy of themselves, and could therefore be seen as instigating a causal chain of events. In this way the use of ill-defined notions prevents an accurate description of the causal meanings of these prepositions.

In general, the grammars use slightly different terms for what seems to be the same notion (such as "instrument", "means", and "instrumental cause"), or use the same term in different ways (such as whether or not inanimates can be "agents"). Some studies that look more specifically at the causal meanings of *bə* and *min* provide better definitions, but the confusion over these terms remains considerable. For instance, Bicknell (1984: 46) defines Agents as "actors or sources of action" (and argues that animacy is a corollary of this definition). Instruments are "inanimate objects with which the action is carried out". But the use of inherent properties (animacy) of the argument instead of relational properties (like whether the argument volitionally participates in the event; cf. Næss 2007: 30–32) to define at least the

that *min* expresses "*from whom* the action comes, *who* is the cause of it", whereas *bə* marks an "instrumental cause" (Joüon & Muraoka 2006: §132de, emphasis original). This chapter can be seen as an attempt to capitalize on that intuition in a more rigorous way. Other grammars do not go into as much detail. Gesenius (1910: §119f) makes a distinction similar to the one of Joüon & Muraoka (2006) when he compares the "*min* of origin" to the "*beth instrumenti*", but provides no definitions of the terms and very few examples. Van der Merwe et al. (2017: §39.1.3.3b) mention that passive Agents can be marked by *bə* and *min*, but in the discussion of these prepositions this function does not return (§39.6.3a, on *bə*, only mentions *bə-yaḏ* 'by the hand of' to refer to Agents; §39.14.4a on *min* mentions an Instrumental function, but the example is not convincing). *Bə* can mark the instrument, cause, or ground, and *min* can mark an instrument or ground (Van der Merwe et al. 2017: §39.11.3ab, 39.14.4b). But this discussion does not offer much over that of Waltke & O'Connor (1990) and Joüon & Muraoka (2006) since it does not define its terms or compare the two prepositions.

Instrument role leads to a number of unusual decisions: for (2.7), Bicknell argues that *bə* marks an Agent because the argument is animate, while most grammarians see it as an Instrument because God is exacting punishment using the man; in (2.6), Bicknell sees the noise as an Instrument, but there is no Agent who uses the noise to split the earth. Other authors have by and large refrained from giving a formal definition of terms like Agent, and have instead proposed tests with which an Agent can be recognized. In this respect one often finds the test that Agents are subjects of prototypical active transitive clauses (e.g. Sollamo 2003; Jones 2018). However, as many of the examples in this chapter show, the causal use of *bə* and *min* is not limited to the passive voice, and there is no reason to think the prepositions behave differently in passive sentences.

2.3 New lexical semantics for *min* and *bə*

The previous section has shown that a traditional approach using thematic roles is only useful to a limited extent to describe the causal meanings of *min* and *bə*. Furthermore, it has little explanatory value, as it does not predict differences in meaning between the two. This is not entirely surprising, since there have been long-standing doubts on the theoretical status of thematic roles like Agent and Instrument.[8] We propose that it is more fruitful to derive the causal meaning of these prepositions from their spatial meanings, using an abstract spatial model of causation.[9] The assumption, based on much cognitive linguistic work (e.g. Radden 1985; Talmy 1988; Dirven 1995;

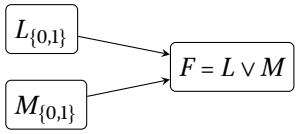
8 For example, it has long been recognized that thematic roles like Cause, Reason, Instrument, and Agent cannot be clearly separated (Dowty 1991; Davis 2011; Harley 2011; Primus 2016). There is no definitive list of such roles, and the boundaries between them are often vague. There has been work on distinguishing Agents and Causers (e.g. Pylkkänen 2008; Martin & Schäfer 2013), but the reference works we are dealing with here use these terms in a less well-defined, intuitive fashion.

9 Of course, this does not mean that earlier findings are entirely wrong. As just one example, it remains true that Instruments marked by *bə* are "typically non-living" (Van der Merwe et al. 2017: §39.6.3a). This is, however, not the most efficient description of the linguistic system. As we see it, inanimacy is an epiphenomenon: an inanimate tool will not be marked by *min* in regular language use, because a tool is always manipulated by another entity and is therefore never fully dominant. We suggest that the description of the causal meanings of *min* and *bə* can be made much simpler and at the same time more precise by referring to the notion of "dominance".

Croft 2012), is that humans conceptualize of causation as a kind of abstract space. Spatial prepositions can be used to express relations in this space. When speakers do this, the prepositions develop a causal meaning based on their spatial meaning. This approach not only circumvents terminological issues with thematic roles, but is also more economical, as it can derive differences in causal meaning without having to store extra information in the lexicon (cf. the notion of principled polysemy developed by Tyler & Evans 2003).

The abstract model that we will employ is that of a causal model (Pearl 2000; Halpern & Pearl 2005, among others). A causal model describes the dependencies between a set of variables. It can be represented as a directed graph, as in (2.15). This model expresses that the occurrence of Fire (F) depends on whether there is lightning (L) and whether a match is lit (M). In this case, all variables range over truth values (indicated by the subscript $\{0,1\}$), and F depends on L and M via inclusive or.

(2.15)



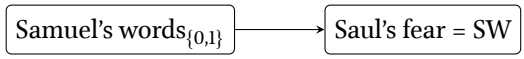
(Halpern & Pearl 2005: 848–850)

In this model, L and M do not depend on other variables. These variables are called exogenous. F is an endogenous variable, as it does depend on other variables.

Now consider (2.1–2.4) again. We repeat only translations here, and provide a causal model for each. The models with *min* are relatively simple:

(2.2′) 1 Sam. 28:20: 'Then Saul fell at once full length on the ground, filled with fear *because of* (*min*) the words of Samuel.' (ESV)

(2.16)



(2.3′) 2 Sam. 7:29: 'Now therefore may it please you to bless the house of your servant, … , and *with* (*min*) your blessing shall the house of your servant be blessed forever.' (ESV)

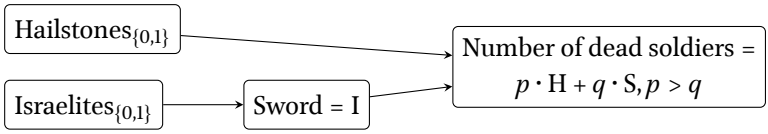
(2.17)

The models described with *bə* are more complex. In (2.18), the number of dead soldiers ranges over non-negative integers rather than truth values, and the formula computing the number of dead soldiers is such that the

hailstones have a greater effect. In (2.19), the land perishes precisely if there is a famine and there are no reserves:

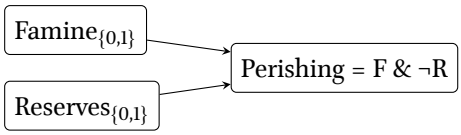
(2.1′) Josh. 10:11: 'There were more who died *because of* (*bə*) the hailstones than the sons of Israel killed *with* (*bə*) the sword.' (ESV)

(2.18)



(2.4′) Gen. 41:36: '(That food shall be a reserve …), so that the land may not perish *through* (*bə*) the famine.' (ESV)

(2.19)



There are two crucial differences between the *min*-causers in (2.16–2.17) and the *bə*-causers in (2.18–2.19). First, causes marked by *min* are represented by exogenous variables: they do not depend on other variables (contrast [2.18], in which the sword marked by *bə* is itself dependent on the action of a volitional Agent). Second, causes marked by *min* are also the *only* exogenous variable on which the effect depends (contrast [2.19], in which the perishing of the land depends not only on the famine marked by *bə*, but on the land's reserves as well). We therefore define dominance as follows:

(2.20) A cause C of an effect E is represented as "dominant" if (a) C is exogenous (not dependent on other variables) and (b) E does not depend on any other exogenous variables besides C.

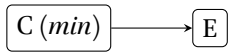
This entails that if there is a dominant cause, the effect is presented as fully determined by that cause. It thus formalizes an intuitive notion of dominance. First, being represented by an exogenous variable, a dominant cause is not caused itself, but rather influences the endogenous variables in the model. Thus, an Agent would be dominant, but an Instrument would not be, since the Instrument variable depends on the Agent variable. Second, a dominant cause precludes the existence of other causes of this type (being the only exogenous variable in the model).

Our claim is that *min* marks such dominant causes, whereas *bə* is unmarked and can in principle mark any cause. However, we can expect *bə* to be blocked from marking dominant causes by Gricean maxims (Grice

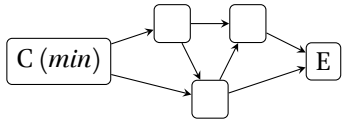
1989): a speaker is required to use the more specific *min* if possible, so that *bə* becomes infelicitous for dominant causes despite its general lexical semantics.[10]

Schematically, the simplest causal model with a *min*-causer is as shown in (2.21a). Example (2.21b) is a more complex model with a *min*-causer: intermediate variables are allowed, as long as the *min*-causer is the only exogenous variable. Two types of causal models that require *bə* are shown in (2.22): in (2.22a), there is more than one exogenous variable; in (2.22b), the causer is not exogenous itself.

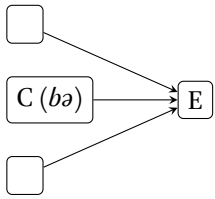
(2.21) a.



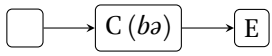
b.



(2.22) a.



b.



Thus, by using *min*, a speaker can make explicit that a causer is (a) not caused itself and (b) ultimately the only factor on which the outcome depends.

This hypothesis trivially derives the well-known fact that *bə* is the default preposition to mark instruments. In (2.23), it is the man who is dominant, and not the rod. The event is represented as in (2.22b), in which the rod must be marked by *bə* and the man is the dominant causer. Similarly, in (2.24), God is invoked by another Agent, and is therefore not a dominant cause for the swearing.

10 An anonymous reviewer suggests that *min* may select '*the* cause' of an event as opposed to '*a* cause' (cf. the discussion of causal selection in Bar-Asher Siegal & Boneh 2020: 21–28). This notion may turn out to overlap with dominance. However, as Bar-Asher Siegal & Boneh (2020: 21) note, the identification of 'the cause' is highly context-dependent. A discussion of the underlying factors would take us too far afield here.

(2.23) Exod. 21:20: וְכִי־יַכֶּה אִישׁ אֶת־עַבְדּוֹ אוֹ אֶת־אֲמָתוֹ בַּשֵּׁבֶט

wə=k̲ī *y-akkɛ-Ø* *ʾīš* *ʾɛt̲* *ʿab̲d-ō* *ʾō* *ʾɛt̲* *ʾămå̄t̲-ō*
and=COMP 3M-strike\IPFV-SG man OBJ servant-his or OBJ maid-his

b=*aš=šēb̲ɛṭ*
in=the=rod

'When a man strikes his slave, male or female, *with* a rod, …' (ESV)

(2.24) Gen. 21:23: וְעַתָּה הִשָּׁבְעָה לִּי בֵּאלֹהִים הֵנָּה

wə=ʿattå̄ *hiššå̄b̲əʿ-å̄* *l-ī* ***b̲***=*ēʾlōhīm* *hēnnå̄*
and=now swear\IMP-M.SG to-me **in**=God here

'Now therefore swear to me here *by* God that …' (ESV)

On the other hand, *min* is often used to describe that something is made impossible (2.25), and *bə* is never used this way. This is an instance of the model in (2.21a), where the *min*-causer alone influences the ability to do something.

(2.25) Gen. 16:10 (also 32:13): הַרְבָּה אַרְבֶּה אֶת־זַרְעֵךְ וְלֹא יִסָּפֵר מֵרֹב׃

harbå̄ *ʾ-arbɛ* *ʾɛt̲* *zarʿ-ēk̲* *wə=lōʾ*
multiply\INFABS 1SG-multiply\IPFV OBJ seed(M)-yours and=not

y-issåp̄ēr-Ø ***mē***=*rōb̲*
3M-count\MID.IPFV-SG **from**=multitude

'I will surely multiply your offspring so that they cannot be numbered *for* multitude.' (ESV)

The proposed lexical semantics of the causal meaning of *min* and *bə* can be derived from the original spatial meaning of these prepositions. Consider the model in (2.26), which has besides the effect *E* both a dominant causer (C_1) and a non-dominant causer (C_2):

(2.26) 

The dependency chain represented by this model can be seen as originating from the dominant causer C_1. It is therefore natural for the dominant causer to be marked by *min*, a preposition that also in the spatial domain marks an origin (as in *miy-yiśrāʾēl* 'from Israel'). Furthermore, the dominant cause is further removed from the effect than the intermediate cause, so it is expected that the dominant cause is marked by a preposition that expresses a greater distance between Figure and Ground than the preposition for the intermediate cause (*min* 'from' as opposed to *bə* 'in'). The causal meanings

of *min* and *bə* can therefore be derived from a mapping of the causal domain onto the spatial domain, consistent with much work in cognitive linguistics (Radden 1985; Talmy 1988; Dirven 1995; Croft 2012, among others).[11]

The behavior of these prepositions aligns particularly well with that described by Croft (2012: 222–226). Croft observed that cross-linguistically causers are typically marked by prepositions with an ablative, perlative, or locative meaning ('from', 'through', and 'in, by, with', respectively), describing locations either close to the Ground ('in', 'by', 'with') or on a path from the Ground to the Figure ('from', 'through'). The prepositions *min* 'from' and *bə* 'in' fit this pattern. What we add to this analysis is the proposal that the causal meaning of a preposition is not only determined by the relative location to the Ground it describes, but also by the distance to the Ground that it expresses: concrete distance in the spatial sense corresponds to abstract distance in the causal model.

2.4 Biblical Hebrew evidence

To test our hypothesis that *min* is used to mark dominant causers, while *bə* is unmarked, we compiled a dataset of causal uses of these two prepositions. Instances were classified as causal when they marked an argument which brings about or plays a facilitating role in the realization of an event or state, in the sense that the effect would not have occurred without intervention of that argument (for this definition, cf. Mackie 1965; Lewis 1973a,b; and see Bar-Asher Siegal & Boneh 2020: 5 for more references). For causers marked by *bə* we relied on the comprehensive overview by Jenni (1992), drawing our examples from rubrics 16 through 19 (*beth causae*, *instrumenti*, *pretii*, and *communicationis*).[12] Causers marked by *min* were collected manually.

We performed a comprehensive analysis of the instances in a number of narrative books, since we expect most unambiguously interpretable causal uses in these texts.[13] We do not have space to discuss every instance here,

11 As already mentioned in chapter 1, footnote 30, it is possible that different languages have different spatial representations of the same causal dependencies. I return to this issue in chapter 6.

12 Not all cases compiled by Jenni are relevant here, but based on the description of the other rubrics these four rubrics form a superset of the data that we are interested in.

13 In particular: Genesis, Exodus, Joshua, Samuel, Ruth, Ecclesiastes, and Esther. These books were chosen to obtain texts from a variety of subgenres and topics. We did not ob-

but our analysis of other instances can be looked up in the data set accompanying this chapter (Staps 2023a). After the initial compilation we used specific verbs and prepositional objects in our dataset to search in the entire Hebrew Bible for more examples for comparison. We excluded examples with a reasonable spatial, temporal, or other non-causal reading of the preposition, even if causation is still implied or contextually inferred. In such cases the preposition may be chosen for reasons other than its causal meaning.[14] We also excluded possibly fossilized uses of *min* and *bə* in complex prepositions (e.g. *mil-lip̄nê* 'from the presence of'), because their meaning may have developed independently (see e.g. Rodriguez 2017; Hardy 2022).

This method uncovered a number of minimal pairs where the same type of event can be caused by both a *min*-causer and a *bə*-causer. As can be seen from the data data set (Staps 2023a), the main evidence suitable for comparison comes from mental states and events. We begin our discussion in section 2.4.1 with a number of relatively straightforward examples concerning the mental state of drunkenness. In section 2.4.2 we discuss mental states expressed by psychological verbs such as *fear* or *be happy*; we finish in section 2.4.3 with mental acts related to volitionality.

2.4.1 Drunkenness

In our corpus, a great number of examples describe someone's mental state after alcohol consumption. With *min* we have the expression *miyyayin* 'from wine':[15]

(2.27) Isa. 51:21: לָכֵן שִׁמְעִי־נָא זֹאת עֲנִיָּה וּשְׁכֻרַת וְלֹא מִיָּיִן׃

lāḵēn	*šimʿ-ī*	*nāʾ*	*zōʾṯ*	*ʿăniy-yāʰ*
therefore	hear\IMP-2F.SG	please	this	afflicted-F.SG

ū=šəḵur-aṯ		*wə=lōʾ*	***miy***=*yāʰyin*
and=become_drunk\PTCP.PASS-F.SG		and=not	**from**=wine

serve differences in causal use between Early and Late Biblical Hebrew, so we consider the corpus homogeneous for the purposes of this chapter.

14 For example, Gen. 17:16 'I will give you a son by (*min* 'from') her' (ESV); 19:32 'let us make our father drink wine, and we will lie with him, that we may preserve offspring from (*min*) our father' (ESV). The preposition *min* can be said to mark a causal relationship here, but it can also be seen as spatial, because the offspring (זֶרַע *zɛra*ʿ 'seed') literally comes out of the father. We have tried to err on the side of caution by selecting the most unambiguously causal examples here.

15 Hos. 7:5 can be explained similarly, but is excluded here because of text-critical issues.

'Therefore hear this, you who are afflicted, who are drunk, but not *with* wine' (ESV)

(2.28) Ps. 78:65: וַיִּקַץ כְּיָשֵׁן אֲדֹנָי כְּגִבּוֹר מִתְרוֹנֵן מִיָּיִן׃

way-y-iqaṣ-Ø	*kə=yǎ̊šēn*	*ʾăḏōn-ǎ̊y*	*kə=ḡibbōr*
WAYQ-3M-awake-SG	like=sleeping	lord-mine	like=strong_man

m-iṯ-rōnēn-Ø	***miy**=yǎ̊yin*
PTCP-REFL-shout\PLURACT-M.SG	**from**=wine

'Then the Lord awoke as from sleep, like a strong man shouting *because of* wine.' (ESV)

With the preposition *bə* we find the expression *ṭōḇ lēḇ* NP *bayyayin* 'NP is good-hearted with wine':

(2.29) 2 Sam. 13:28: רְאוּ נָא כְּטוֹב לֵב־אַמְנוֹן בַּיַּיִן וְאָמַרְתִּי אֲלֵיכֶם הַכּוּ אֶת־אַמְנוֹן וַהֲמִתֶּם אֹתוֹ

rəʾ-ū	*nǎ̊ʾ*	*kə=ṭōḇ-Ø*	*lēḇ*	*ʾamnōn*
watch\IMP-M.PL	please	like=good-M.SG.of	heart	Amnon

***b**=ay=yayin*	*wə-ʾǎ̊mar-tī*	*ʾălē-ḵɛm*	*hak-kū*	*ʾɛṯ*
in=the=wine	WQAT-say-1SG	to-you	strike\IMP-M.PL	OBJ

ʾamnōn	*wa-hămit-tɛm*	*ʾōṯ-ō*
Amnon	WQAT-die\CAUS-2M.PL	OBJ-him

'Mark when Amnon's heart is merry *with* wine, and when I say to you, "Strike Amnon," then kill him.' (ESV)

(2.30) Est. 1:10–11: בַּיּוֹם הַשְּׁבִיעִי כְּטוֹב לֵב־הַמֶּלֶךְ בַּיָּיִן אָמַר ... לְהָבִיא אֶת־וַשְׁתִּי הַמַּלְכָּה לִפְנֵי הַמֶּלֶךְ בְּכֶתֶר מַלְכוּת

b=ay=yōm	*haš=šəḇīʿī*	*kə=ṭōḇ-Ø*	*lēḇ*	*ham=mɛlɛḵ*
in=the=day	the=seventh	like=good-M.SG.of	heart	the=king

***b**=ay=yǎ̊yin*	*ʾǎ̊mar-Ø*	*...*	*lə=hǎ̊ḇīʾ*	*ʾɛṯ*	*waštī*
in=the=wine	say\PFV-3M.SG	...	to=bring\INF	OBJ	Vashti

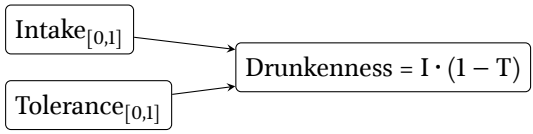
ham=malkǎ̊	*lipnē*	*ham=mɛlɛḵ*	*bə=ḵɛṯɛr*	*malḵūṯ*
the=queen	before	the=king	in=crown.of	kingship

'On the seventh day, when the heart of the king was merry *with* wine, he commanded ... to bring Queen Vashti before the king with her royal crown' (ESV)

Suppose that we measure drunkenness on a scale from sober to fully intoxicated, represented in a causal model by a real-valued variable. In the

model in (2.31), a person's drunkenness is modeled as dependent on both alcohol intake and someone's personal alcohol tolerance:[16]

(2.31)



In (2.29–2.30), the drinker's alcohol tolerance still reduces the level of drunkenness. In (2.30), the king is still capable of giving commands and engaging in a discussion of law (Est. 1:15–20). In (2.29) it is likewise not required that Amnon is knocked out by the drink; he only needs to be off his guard. In both cases, then, the degree of drunkenness is not fully dependent on the alcohol intake. The preposition *bə* is used because alcohol intake is not the only exogenous variable in the model.

This is different in (2.27–2.28). The use of *min* in these examples excludes other exogenous variables from the model: the alcohol intake is so high that no level of alcohol tolerance can reduce the degree of drunkenness. In (2.27), 'drunk with (*min*) wine' is used as a description of the inhabitants of Jerusalem, who are afflicted by 'devastation and destruction, famine and sword' (v. 19). The high degree of drunkenness implied by the use of *min* is used as a metaphor for this severe affliction. Example (2.28) is to be understood similarly. The psalm goes on to describe how God puts his adversaries to everlasting shame. The frightening image of 'a strong man' completely inebriated by alcohol is used to indicate the kind of fear the adversaries should have on account of God. By removing possible tempering causers (someone's alcohol tolerance) from the model, the use of *min* implies that the effect of the cause is severe.

The following example with both *bə* and *min* can be used as further evidence:

(2.32) Isa. 28:7: וְגַם־אֵ֣לֶּה בַּיַּ֣יִן שָׁגׇ֔וּ וּבַשֵּׁכָ֖ר תָּעׇ֑וּ כֹּהֵ֤ן וְנָבִיא֙ שָׁגׇ֣וּ בַשֵּׁכָ֔ר נִבְלְעׇ֖וּ מִן־הַיַּ֑יִן תָּעׇ֣וּ מִן־הַשֵּׁכָ֔ר

wə=gām	*ʾēllɛ*	***b**=ay=yayin*	*šɔ̄g-ū*	
and=also	these	**in**=the=wine	go_astray\PFV-3PL	
*ū=**b**=aš=šēḵɔ̄r*		*tɔ̄ʿ-ū*	*kōhēn*	*wə=nɔ̄bīʾ*
and=**in**=the=strong_drink		stagger\PFV-3PL	priest	and=prophet

16 In this model, the subscript [0, 1] indicates that variables range over real values between 0 and 1 (inclusive).

šå̄ḡ-ū	***b̲***=*aš*=*šēk̲å̄r*	*niḇləʿ-ū*	***min***
go_astray\PFV-3PL	**in**=the=strong_drink	numb\MID.PFV-3PL	**from**

hay=*yayin*	*tå̄ʿ-ū*	***min***	*haš*=*šēk̲å̄r*
the=wine	stagger\PFV-3PL	**from**	the=strong_drink

'These also reel *with* wine and stagger *with* strong drink; the priest and the prophet reel *with* strong drink, they are swallowed/confused *by* wine, they stagger *with* strong drink' (ESV)

The writer has combined verbs and prepositional objects to produce a climactic sequence:

1. *Šå̄ḡå̄* ***b̲***ayyayin 'go astray', '*in*', 'the wine'
2. *Tå̄ʿå̄* ***b̲***aššēk̲å̄r 'stagger', '*in*', 'the strong drink'
3. *Šå̄ḡå̄* ***b̲***aššēk̲å̄r 'go astray', '*in*', 'the strong drink'
4. *Niḇlaʿ* ***min*** *hayyayin* 'be numbed', '*from*', 'the wine'
5. *Tå̄ʿå̄* ***min*** *haššēk̲å̄r* 'stagger', '*from*', 'the strong drink'

Taken separately, the verbs are not in strictly ascending order in terms of severity, and *yayin* and *šēk̲å̄r* are not strictly ordered by strength either.[17] This is presumably done to avoid a highly repetitive pattern. When the verbs and nouns are taken together, it is clear that the text is climactic. This climax is also mirrored in the choice of prepositions (three times *bə* followed by two times *min*). This verse therefore lends further support for the claim that *min* marks dominant causers.

2.4.2 Psychological verbs

Psychological verbs describe the mental state of an Experiencer, often with respect to some other argument. This mental state can be expressed in one of two ways: (a) as a two-place relation between the Experiencer and a Target or Subject Matter, as in *Sue delights in the rain*, or (b) as a one-place property of the Experiencer, which may optionally be caused by a Cause, as in *Sue is delighted* (*because of the rain*) (Pesetsky 1995; Doron 2020; among

17 We take *šēk̲å̄r* 'strong drink' to be stronger than *yayin* 'wine'. Although the two words often stand parallel to each other, the root *škr* more frequently has pejorative overtones (piel/hiphil *škr* 'make lose control'; *šikkōr* 'utterly drunk' in e.g. 1 Sam. 1:13); cf. Oeming (2006: 1–2). The text thus contains two climactic sequences if we look at the nouns (items 1–3 and 4–5). For the verbs, we assume based on other uses that *tʿh* 'stagger' is stronger than *šgh* 'go astray', with *nblʿ* 'be numbed' somewhere in between. The verbs are thus also ordered in two climactic sequences (items 1–2 and 3–5).

others). According to Doron (2020: 409–410), prepositions used to mark the Target/Subject Matter with relational verbs are varied, but lexically selected by the verb (e.g., *rejoice*/*revel*/*delight* + *in* vs. *enjoy*/*like* + direct object). On the other hand, the prepositions used to mark the Cause with property verbs are chosen from a small set of "causal prepositions": prepositions that have causal uses independent of psychological verbs, like *because of* and *due to*.

Property verbs are most relevant to us here, since both *min* and *bə* can be used with these verbs and then have their general causal meaning. However, relational verbs that happen to take both *min* and *bə* are relevant as well, since we would expect the meanings of these prepositions with the given verb to still reflect their general causal meanings, even if the verb-preposition pair is partially lexicalized. For both types of psychological predicates, we therefore argue that the choice of preposition makes a difference in interpretation, according to the hypothesis spelled out in section 2.3.[18]

2.4.2.1 Fear

We find *min* and *bə* marking the cause of a variety of mental states expressed by psychological verbs. The tendency is for causes of negative mental states to be marked by *min*, while causes for positive mental states are marked by *bə*. As we will explain below, this follows from the fact that *min* marks dominant causers. Furthermore, we will show that the exceptions to this general pattern can be explained by the notion of dominance as well.

18 Psychological verbs often appear in doublets, where the Experiencer is the subject with one verb and the object with the other (e.g. *x likes y* vs. *y pleases x*; cf. Dowty 1991: 579). Though most of the verbs we discuss here have such a doublet, this feature is not relevant for our discussion, so we focus here on any verb that expresses a mental state with respect to some Target/Subject Matter (relational verbs) or brought about by a Cause (property verbs). Note that some roots can be used as both a relational and a property verb. According to Doron (2020), in Modern Hebrew, relational verbs appear in the intensive (pluractional) template when the Experiencer is the object, while property verbs appear in the causative template in this case. This seems to be the case in Biblical Hebrew as well. Importantly, some roots appear with both templates, and then have the expected difference in meaning (e.g. *śmḥ* 'rejoice': relational 2 Chr. 20:27 'the Lord has made them rejoice [piel, intensive/pluractional] *over their enemies*'; property Ps. 89:43 'you have made all his enemies rejoice [hiphil, causative]'). To determine whether a verb is used to describe a relation or a property, one must look at the meaning in context (whether it is relational or not) and, sometimes, preposition (a verb with a non-causative preposition is necessarily relational). However, the classification as a relational or property verb is, as explained in the main text, not crucial to our argument.

One of the most frequent examples of negative mental states caused by a *min*-causer is the state of fear:[19]

(2.33) Exod. 3:6:[20] וַיַּסְתֵּר מֹשֶׁה פָּנָיו כִּי יָרֵא מֵהַבִּיט אֶל־הָאֱלֹהִים׃

way-y-astēr-Ø *mōšɛ* *på̄n-å̄yw* *kī* *yårē'-Ø*
WAYQ-3M-hide-SG Moses face-his COMP fear\PFV-3M.SG

***mē**=habbīṭ* *'ɛl* *hå̄='ĕlōhīm*
from=look\INF to the=God

'And Moses hid his face, for he was afraid *to* look at God.' (ESV)

(2.34) Eccl. 12:5: גַּם מִגָּבֹהַּ יִרָּאוּ וְחִתְחִתִּים בַּדָּרֶךְ

gam ***mig**=gå̄ḇōah* *y-irå̄'-ū* *wə=ḥaṯḥatt-īm* *b=ad=dɛrɛḵ*
also **from**=high 3M-fear\IPFV-PL and=terror-PL in=the=way

'they are afraid also[21] *of* what is high, and terrors are in the way' (ESV)

Since fear is typically something that *overcomes* an Experiencer, *min* fits these contexts well: the reason for fear (e.g., the thought of looking at God in [2.33]) overcomes the Experiencer and thereby excludes any other possible influences on their mental state, excluding causal models such as (2.22a).

Unsurprisingly, Jenni (1992: 112–113) finds only one case where a cause for fear is marked by *bə*:

(2.35) Jer. 51:46: וּפֶן־יֵרַךְ לְבַבְכֶם וְתִירְאוּ בַּשְּׁמוּעָה הַנִּשְׁמַעַת בָּאָרֶץ

ū=pɛn *y-ēraḵ-Ø* *ləḇaḇ-ḵɛm* *wə=ṯ-īrə'-ū*
and=lest 3M-be_weak\IPFV-SG heart(M)-yours and=2M-fear\IPFV-PL

***b**=aš=šəmūʿå̄* *han=nišmaʿ-aṯ* *b=å̄='årɛṣ*
in=the=report(F) the=hear\MID.PTCP-F.SG in=the=land

'Let not your heart faint, and be not fearful // *at* the report heard in the land' (ESV)

We prefer to read this instance of *bə* temporally: '*when* the report is heard in the land'.

19 Consider also, without *yr'* 'fear', 1 Sam. 1:16: 'I have been speaking out of (*min*) my great anxiety (*śīaḥ*) and vexation (*kaʿas*)' (ESV).

20 Similarly Exod. 34:30.

21 Here, *gam* 'also' does not imply that 'what is high' is but one of several reasons for fear. Though it is difficult to interpret the structure of the poem, commentators are generally agreed that *gam* applies to the entire clause and describes a way in which the 'days are bad' (12:1) in addition to the other ways mentioned in 12:1–7 (e.g. Krüger 2004: 200–201). We should thus read this as *gam* [*miggāḇōah yīrā'û*] rather than [*gam miggāḇōah*] *yīrā'û*, so the cause for fear is dominant here.

2.4.2.2 Shame

Also causes for shame are almost exclusively marked by *min*:[22]

(2.36) Jer. 22:22: כִּי אָז תֵּבֹשִׁי וְנִכְלַמְתְּ מִכֹּל רָעָתֵךְ׃

kī	*ʾāz*	*t-ēḇōš-ī*	*wə-niḵlam-t*
COMP	then	2-be_ashamed\IPFV-F.SG	WQAT-be_ashamed-2F.SG

mik=*kōl*	*rāʿāṯ-ēḵ*
from=all.of	evilness-yours

'then you will be ashamed and confounded *because of* all your evil.' (ESV)

Shame may be a more subtle emotion than fear, so it may not be immediately clear that a reason for shame can overcome an Experiencer. How then do we explain the use of *min*? It may be relevant that almost all of the occurrences are in the prophetic literature. In this genre, the reason for shame is typically presented as utterly humiliating. In that sense, it can be seen as excluding other potential influences on the Experiencer's mental state (as we saw with fear), which may explain the preference for *min*. It is noteworthy that the only contrasting example with *bə* comes from outside the prophetic corpus:[23]

(2.37) Ps. 69:7: אַל־יֵבֹשׁוּ בִּי קֹוֶיךָ אֲדֹנָי יְהוִה צְבָאוֹת אַל־יִכָּלְמוּ בִי מְבַקְשֶׁיךָ

ʾal	*y-ēḇōš-ū*	***ḇ***-*ī*	*qōw-ɛ-ḵā*	*ʾăḏōn-āy*
not	3M-be_ashamed\JUSS-PL	**in**-me	trust\PTCP-M.PL-yours	lord-mine

yhwh	*ṣəḇāʾ-ōṯ*	*ʾal*	*y-ikkāləm-ū*	***ḇ***-*ī*
Yahweh.of	host-PL	not	3M-be_ashamed\JUSS-PL	**in**-me

m-əḇaqš-ɛ-ḵā
PTCP-seek-M.PL-yours

'Let not those who hope in you be put to shame *through* me, O Lord God of hosts; let not those who seek you be brought to dishonor *through* me' (ESV)

22 We are grateful to an anonymous reviewer for suggesting these cases. Jenni (1992: 112) only mentions one example with *bə*, with the roots *klm* (niphal) and *bûš* (qal), both meaning 'to be ashamed' (example [2.37]). We then looked for examples with these roots and the preposition *min*. With niphal *klm* 'be ashamed', see further Ezek. 16:27, 54; 36:32; 43:10, 11; with *bûš* 'be ashamed', Isa. 1:29; 20:5; Jer. 2:36; 12:13; 48:13; Ezek. 32:30; Hos. 4:19, 10:6; Mic. 7:16; Zeph. 3:11; Zech. 13:4; Ps. 119:116. In Isa. 1:29 we also find *ḥpr* 'be ashamed' with *min*.

23 There are two more possible cases in 2 Sam. 19:4 and Ezek. 16:61, but they are easier read with a temporal sense.

Besides the difference in genre, it may also be relevant that this instance is found in a negated wish. It has a humbling effect: the psalmist does not want to cause shame in any way, whether dominantly so or not. We return to the interpretation of causal *min* and *bə* in negated sentences in section 2.5 below.

2.4.2.3 Joy

Of the positive mental states, joy is the most frequent. As already mentioned above, causes for positive mental states tend to be marked by *bə*. Jenni (1992: 106–108) lists 91 cases of causes for joy marked by *bə*, predominantly with *śmḥ* 'rejoice', *gyl* 'shout out', and hithpael *hll* 'boast':

(2.38) Jdg. 9:19: שִׂמְחוּ בַּאֲבִימֶלֶךְ וְיִשְׂמַח גַּם־הוּא בָּכֶם׃

śimḥ-ū	***ba***=*ʾăḇīmɛlɛḵ*	*wə=y-iśmaḥ-Ø*	*gam*	*hū'*
rejoice\IMP-M.PL	**in**=Abimelech	and=3M-rejoice\JUSS-SG	also	he

bå̊-*ḵɛm*
in-you

'rejoice *in* Abimelech, and let him also rejoice *in* you.' (ESV)

(2.39) Eccl. 5:19: וְלִשְׂמֹחַ בַּעֲמָלוֹ זֹאת מַתַּת אֱלֹהִים הִיא׃

wə=li=śmōaḥ	***ba***=*ʿămå̊l-ō*	*zōh*	*matta-ṯ*	*ʾĕlōhīm*	*hī'*
and=to=rejoice\INF	**in**=toil-his	this	gift-of	God	it

'(God has given man the ability ... to accept his lot) and to rejoice *in* his toil — this is the gift of God.' (ESV)

In (2.38), the Addressee has to be told to 'rejoice in Abimelech'. This shows that Abimelech alone is not sufficient cause for the joy of the Addressee: in addition to the presence of Abimelech, a command is needed, making the underlying causal model a variant of (2.22a). The Addressee needs to be actively involved in generating that joy. The same is true for (2.39), since 'toil', within the context of Ecclesiastes, is not something that sparks joy in and of itself. In this sense, these causers are not dominant, since the Experiencer's active participation is another factor in the causal model. This explains why these causers are not marked by *min* but by *bə*.

A rare exception to the overall tendency to use *bə* shows that this is indeed how the preposition should be understood:

(2.40) Prov. 5:18: יְהִי־מְקוֹרֶךָ בָּרוּךְ וְשְׂמַח מֵאֵשֶׁת נְעוּרֶיךָ׃

y-əhī-Ø	*məqōr-əḵå̄*	*b̲å̄rūḵ-Ø*
3M-be\JUSS-SG	fountain(M)-yours	bless\PTCP.PASS-M.SG
ū=śəmaḥ-Ø	***mē***=*ʾēšɛṯ*	*nəʿūrɛ-ḵå̄*
and=rejoice\IMP-M.SG	**from**=wife.of	youth-yours

'Let your fountain be blessed, and rejoice *in* the wife of your youth' (ESV)

This verse appears in a chapter full of warnings against adultery, so it should be understood not only as an instruction to 'rejoice in the wife of your youth' (as opposed to doing something entirely different), but first and foremost as a warning not to rejoice in any other wife. The 'wife of your youth' is a source of joy to the exclusion of other sources of joy (thus precluding a model like [2.22a]); in other words, she is the dominating factor. This explains the use of *min*: the use of this preposition indicates that its argument dominates any other hypothetical reason for joy.

2.4.2.4 Having (had) enough

In this paragraph we discuss the root *śbʿ* 'have (had) enough of something, find something enough'.[24] Though this root is not inherently positive or negative, it does entail that the subject evaluates the object (namely, as being 'enough'), and thus qualifies as a relational psychological verb (for the notion of evaluation in relation to psychological verbs, see Pesetsky 1995: 56). It is most frequent with *min*, but examples with *bə* are not uncommon.[25] With this root, *min* most clearly implies that the causer is the *only* source for satisfaction, as in (2.41) as opposed to (2.42):[26]

(2.41) Eccl. 6:3: וְנַפְשׁוֹ לֹא־תִשְׂבַּע מִן־הַטּוּבָה

wə=napš-ō	*lō'*	*ṯ-iśbaʿ-Ø*	***min***	*haṭ=ṭōb̲å̄*
and=soul-his	not	3F-be_satisfied\IPFV-SG	**from**	the=good

'(If a man fathers a hundred children and lives many years …), but his soul is not satisfied with [life's] good things, …' (ESV)

24 We are grateful to an anonymous reviewer for drawing our attention to this root.

25 Some of the instances of *bə* may be best understood as instrumental: Ps. 103:5; Lam. 3:15, 30. We focus here on the clearest examples.

26 Similar examples with *min* are Ezek. 32:4; Ps. 104:13; Job 19:22; Prov. 1:31; 12:14; 14:14; 18:20. With *bə*: Ps. 65:5. One can also compare the niphal of *ml'* 'be full': with *min* Ezek. 32:6; with *bə* Prov. 24:4. With *rwh* 'be saturated' we only find *min*: Isa. 34:7; Jer. 46:10.

(2.42) Ezek. 16:29: וַתַּרְבֶּי אֶת־תַּזְנוּתֵךְ ... וְגַם־בְּזֹאת לֹא שָׂבֵעַתְּ:

wat-t-arb-ī	*ʾɛṯ*	*taznūṯ-ēk*	...	*wə=gam*	***bə***=*zōʾṯ*	*lōʾ*
WAYQ-2-multiply-F.SG	OBJ	whoring-yours	...	and=also	**in**=this	not

śå̄bå̄ʿ-at
be_satisfied\PFV-2F.SG

'You multiplied your whoring ... and even *with* this you were not satisfied' (ESV)

In (2.41), the use of *min* implies that the writer would expect the man who lives many years to be satisfied from life's good things alone (as opposed to being satisfied from these things in combination with other circumstances; contra a model like [2.22a]; for the interpretation of *min* with negation, see section 2.5 below). On the other hand, in (2.42), the 'whoring' is presented as yet another sin of the Addressee, which even in addition to (*gam* 'even, also') the earlier sins does not satisfy them. The fact that this source of satisfaction is not the *only* source requires the use of *bə*, as it suggests a model like that in (2.22a).

The contribution of *min* can here be shown by the addition of 'alone' in translation: (2.41′) is an accurate translation of (2.41) but (2.42′) does not faithfully reflect (2.42):

(2.41′) 'and his soul is not satisfied with life's good things *alone*'

(2.42′) #'... and even with this *alone* you were not satisfied'

This follows directly from the definition of dominance in (2.20) above.

2.4.2.5 Psychological verbs: conclusion

As has become clear, the general tendency is for psychological verbs describing a negative mental state to have causers marked by *min*, whereas verbs describing a positive mental state tend to have causers marked by *bə*.[27] This

27 If quantitative data is desired, consider the frequency with which *bə* and *min* mark causes for joy and fear, the most common categories in the corpus. Whereas Jenni (1992) lists 91 cases where *bə* marks a cause for joy (primarily in Isaiah and Psalms; 6 in our corpus), we found only three cases with *min* (two in our corpus, in the same verse) by looking at other occurrences of the same verbs and phrases that occur with *bə* (Mannati 1970 suggests that *min* in Ps. 4:8 may be causal as well, but the passage remains troublesome). For causes for fear, the distribution is completely different. Jenni counts only seven cases where *bə* marks the cause of a negative inner process (zero in our corpus), of which one for fear. Here, *min* is more frequent, with ten examples

is also true for verbs with which we find only one preposition. These verbs do not lend themselves for a comparison of minimal pairs and the preposition may be lexically selected by the verb, but nevertheless confirm the hypothesis (with *min*: *dʾg* 'worry' [Josh. 22:24; Jer. 42:16; Ps. 38:19]; *dll* 'be low' [Jdg. 6:6]; with *bə*: hithpael *ḥzq* 'strengthen oneself' [1 Sam. 30:6]). The generalization extends beyond the class of psychological verbs; thus with *min* we also find *ʾnḥ* 'groan' (Exod. 2:23), *blh* 'wear out' (Josh. 9:13), *zʿq* 'cry out' (Exod. 2:23; 1 Sam. 8:18; Isa. 26:17; Job 35:9; 2 Chr. 20:9), and *šwʿ* 'cry out' (Exod. 2:23). The verb *mûṯ* 'die' occurs almost exclusively with *bə*, though many instances can be seen as instrumental or circumstantial. The single causal use of *min* has *mippənē* 'from the face of' (Jer. 38:9). It fits our hypothesis, but we leave open the possibility that *mûṯ* lexically selects *bə* to mark causers as the result of a TEMPORAL > CAUSE shift (Kuteva et al. 2019: 425) and that this one instance with *min* is idiosyncratic. With the exception of *mûṯ* 'die', all frequent verbs occur with both *min* and *bə*.

As we have argued above, the fact that positive mental states co-occur with *bə*, while negative mental states co-occur with *min*, is likely a side effect of the dominance expressed by *min*. At least in our corpus, negative mental states such as fear and shame (in the prophetic literature) are typically felt as things that *overcome* the Experiencer. Causes for these emotions are marked by *min* because they take control and cancel out other possible intervening causers, in much the same way as excessive alcohol intake does (see section 2.4.1). By contrast, the causes for positive mental states like joy do not typically cause a kind of overjoyed mental state that cancels out any other possible intervening causers, so *bə* is a more appropriate preposition for these causers: the corresponding causal model may have more than one exogenous variable, so the *bə*-causer is not dominant.

2.4.3 Intentions and taking action

Another environment in which we find both *min*-causers and *bə*-causers is that of decisions. The distribution is that *min* is used to mark causers who volitionally *take* a decision (2.10; 2.43), whereas *bə* is used for factors influencing such decisions (2.44–2.46):

(2.10) 2 Sam. 3:37: כִּי לֹא הָיְתָה מֵהַמֶּלֶךְ לְהָמִית אֶת־אַבְנֵר בֶּן־נֵר׃

in our corpus alone. Clearly, *bə* is better suited for marking causes for joy and *min* is better suited for marking causes for fear.

kī lōʾ hăyə-ṯă ***mē***=*ham*=*mɛlɛḵ lə*=*hămīṯ ʾɛṯ ʾaḇnēr*
COMP not be\PFV-3F.SG **from**=the=king to=die\CAUS.INF OBJ Abner
bɛn nēr
son.of Ner

'... that it had not been the king's *will* (lit.: *from* the king) to put to death Abner the son of Ner' (ESV)

(2.43) Gen. 24:50: מֵיְהוָה יָצָא הַדָּבָר לֹא נוּכַל דַּבֵּר אֵלֶיךָ רַע אוֹ־טוֹב׃

mē=*yhwh yăṣăʾ-Ø had*=*dăḇăr lōʾ n-ūḵal*
from=Yahweh go_out\PFV-3M.SG the=matter(M) not 1PL-be_able\IPFV
dabbēr ʾēlɛ-ḵă raʿ ʾō ṭōḇ
speak\INF to-you bad or good

'The thing has come *from* the Lord; we cannot speak to you bad or good.' (ESV)

(2.44) Exod. 16:8: בְּתֵת יְהוָה לָכֶם בָּעֶרֶב בָּשָׂר לֶאֱכֹל ... בִּשְׁמֹעַ יְהוָה אֶת־תְּלֻנֹּתֵיכֶם

bə=*ṯēṯ yhwh lă-ḵɛm b*=*ă*=*ʿɛrɛḇ băśăr lɛ*=*ʾĕḵōl ...*
in=give\INF Yahweh to-you in=the=evening meat to=eat\INF ...
bi=*šmōaʿ yhwh ʾɛṯ təlunn-ōṯ-ēḵɛm*
in=hear\INF Yahweh OBJ grumbling-PL-yours

'When the Lord gives you in the evening meat to eat ..., *because* the Lord has heard your grumbling' (ESV)

(2.45) Gen. 15:8: וַיֹּאמַר אֲדֹנָי יֱהוִה בַּמֶּה אֵדַע כִּי אִירָשֶׁנָּה׃

way-y-ōʾmar-Ø ʾăḏōn-ăy yhwh ***bam***=*mă* *ʾ-ēḏaʿ kī*
WAYQ-3M-say-SG lord-mine Yahweh **in**=what 1SG-know\IPFV COMP
ʾ-īrăš-ɛnnă
1SG-inherit\IPFV-it.OBJ

'But he said, "O Lord God, how (lit.: *by* what) am I to know that I shall possess it?"' (ESV)

(2.46) Gen. 42:15: בְּזֹאת תִּבָּחֲנוּ חֵי פַרְעֹה אִם־תֵּצְאוּ מִזֶּה כִּי אִם־בְּבֹוא אֲחִיכֶם הַקָּטֹן הֵנָּה׃

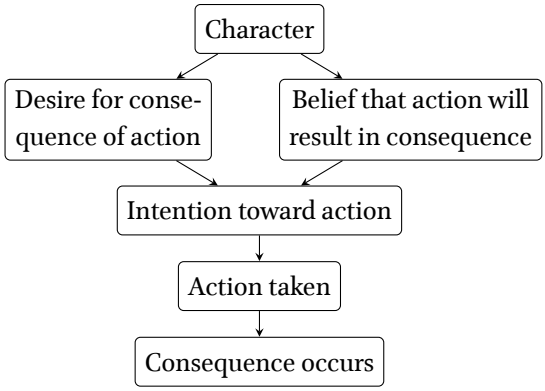
bə=*zōʾṯ t-ibbăḥēn-ū ḥ-ē pārʿōh ʾim t-ēṣəʾ-ū*
in=this 2M-test\MID.IPFV-PL life-of Pharaoh if 2M-go_out\IPFV-PL
miz=*zɛ kī ʾim bə*=*ḇōʾ ʾăḥī-ḵɛm haq*=*qăṭōn hēnnă*
from=this but if in=come\INF brother-yours the=small here

'*By* this you shall be tested: by the life of Pharaoh, you shall not go from this place unless your youngest brother comes here.' (ESV)

In (2.43), Abraham's servant has asked Laban and Bethuel if he may take Rebekah (Laban's sister and Bethuel's daughter) as a wife for Isaac (Abraham's son), since God has pointed out Rebekah to him. Laban and Bethuel answer that it is not their place to question a decision of Yahweh (a 'thing … come from the Lord').[28] The use in (2.10) is similar. On the other hand, in (2.44) *bə* marks the reason why God decides to provide food, and in (2.45–2.46) *bə* marks signs that can be interpreted to obtain knowledge (which can be seen as reasons for deciding to consider something to be the case).

We argue that the distinction between causers *taking* a decision (which are marked by *min*) and causers *influencing* a decision (which are marked by *bə*) boils down to a difference in dominance. In particular, the distribution matches what Malle & Knobe (1997) call the "folk concept of intentionality". In this model, a person's intention depends on (a) a desire to obtain a certain outcome and (b) beliefs about the world concerning how this outcome can be brought about. Whether an action is then taken depends on the person's intention and their skills. Sloman et al. (2012) summarize this with the following causal model:[29]

(2.47)



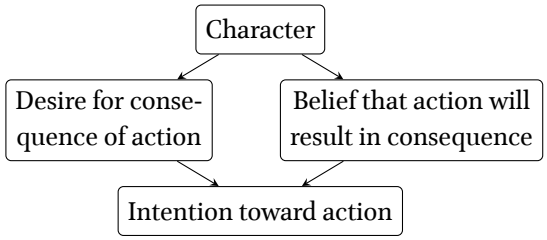
For instance, if someone has the desire to have many flowers in their garden, and believes that watering regularly will bring about that effect, they will have the intention to do so. They will take the action if they also have the skills to perform it, which can then lead to the desired effect.

28 Other decisions made by Yahweh marked by *mē-yhwh* 'from Yahweh' are for Samson to take a Philistine wife (Jdg. 14:4); the kingdom to be given to Solomon (1 Kgs. 2:15); the roll of the dice (Prov. 16:33); and perhaps to assemble Israel (1 Chr. 13:2).

29 For simplicity, some variables irrelevant to our discussion have been left out.

There is an important difference between the examples with *min* (2.10; 2.43) and those with *bə* (2.44–2.46), which can be captured using this model. The examples in (2.10) and (2.43) are primarily about the intention of the Agent (whether the king *wanted* to kill Abner; whether the *decision was made* by God), not whether the event actually occurs. These examples therefore only describe part of the model in (2.47), as in (2.48):

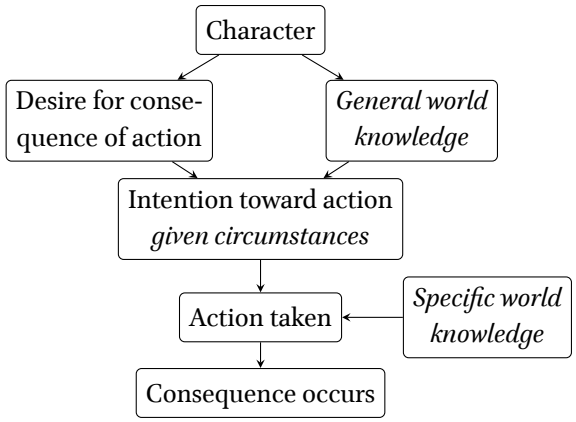
(2.48)



In this model, the volitional Agent ("character") is a dominant causer for their intention, which explains why this dependency is marked by *min* in (2.10) and (2.43).

The sentences in (2.44–2.46) describe a different causal dependency, namely the dependency of the occurrence of the action or consequence (the provision of meat; obtaining knowledge) on something external (hearing the people's grumbling in [2.44] and a sign in [2.45–2.46]). To incorporate this external variable in the causal model in (2.47), we propose the revised model in (2.49). In this revision, the "belief that action will result in consequence" is split between general world knowledge and specific world knowledge (note that in this revision, the character is still a dominant causer for intention, as in [2.48]).

(2.49)



In the example of watering the garden above, general world knowledge would be: "plants need water to grow". The person's intention then becomes: "*if* there is no rain, *then* I intend to water the plants". Whether the action (watering the plants) is executed now depends on specific world knowledge (whether it has rained recently).

Similarly, whether God provides meat in (2.44) depends not only on his intention not to let the people starve but also on the specific world knowledge that there is not enough food. In (2.45) and (2.46), the specific world knowledge includes the sign that can be interpreted by the character to obtain knowledge (the "action"). Crucially, in the model in (2.49), specific world knowledge is not a dominant cause of the action, which also depends on the character's intention: this model is an instance of (2.22a). This explains why these causal dependencies are marked by *bə*.

In sum, passages where decisions are being made describe relatively complex causal models. These models involve the decision-maker, factors influencing their decision, the decision, and possibly the outcome. In an intuitive model, the hypothesis that *min* marks dominant causers explains why decision-makers are marked by *min*, while factors influencing their decision are marked by *bə*.

2.5 Effect under negation

Before concluding, we briefly discuss the effect of negation on the meaning of *min*. Consider first the English verb *stop*. Intuitively, *X stopped verbing* implies not only (a) that *X* does not currently *verb*, but also (b) that *X* previously *verbed* (2.50a). When the sentence is negated, the (a)-implication no longer holds, but the (b)-implication still does (2.50b). The (a)-implication is an at-issue entailment, while the (b)-implication is a presupposition: it is taken for granted (see Kadmon 2001: 10–15 for an introduction).

(2.50) a. *Sue stopped drinking.*
⇒ (a) Sue does not drink; (b) Sue drank.

b. *Sue's problem is that she hasn't stopped drinking.*
⇒ (b) Sue drank. (cf. Kadmon 2001: 10)

Something similar is the case for *min*. Under our definition of dominance in (2.20), this preposition carries two implications: (a) that the object is a causer, and (b) that the object is dominant. In our view, the second

implication is a presupposition, since the dominance of the prepositional object is preserved under negation. Consider again (2.10–2.11):

(2.10) 2 Sam. 3:37: כִּי לֹא הָיְתָה מֵהַמֶּלֶךְ לְהָמִית אֶת־אַבְנֵר בֶּן־נֵר׃

kī *lōʾ* *hăyə-ṯă* ***mē***=*ham*=*mɛlɛḵ* *lə*=*hămīṯ* *ʾɛṯ* *ʾaḇnēr*
COMP not be\PFV-3F.SG **from**=the=king to=die\CAUS.INF OBJ Abner

bɛn *nēr*
son.of Ner

'... that it had not been the king's *will* (lit.: *from* the king) to put to death Abner the son of Ner' (ESV)

(2.11) Gen. 9:11: וְלֹא־יִכָּרֵת כָּל־בָּשָׂר עוֹד מִמֵּי הַמַּבּוּל

wə=*lōʾ* *y-ikkărēṯ-Ø* *kål* *bāśār* *ʿōḏ* ***mim***=*m-ē*
and=not 3M-cut_off\MID.IPFV-SG all.of flesh(M) again **from**=water-of

ham=*mabbūl*
the=flood

'... that never again shall all flesh be cut off *by* the waters of the flood' (ESV)

If the implication that the object of *min* is dominant were a simple entailment, (2.10) would be felicitous if the king were not a dominant causer for Abner's death, but only a non-dominant causer. For example, this sentence would be felicitous if the king had collaborated with others to bring about Abner's death. In context, however, it is clear that any involvement of the king must be excluded. Similarly, in (2.11) the promise is not merely that a flood will never again be the *sole* cause for destruction; rather, it is that a flood will never again be involved in 'cutting off all flesh' in any way. Other cases of negation are similar.[30] Therefore, the dominance of the prepositional object of *min* is preserved under negation, so this aspect of the meaning is presuppositional.[31] It is important to keep this in mind for the correct interpretation of examples with *min*.

30 Other examples involving a *min*-causer under negation are Gen. 46:3; Josh. 10:8; Eccl. 6:3 (example [2.41]); 7:10; Est. 5:9; and outside our corpus Isa. 51:21 (example [2.27]).

31 Recall that we have described *bə* as a more general preposition: both prepositions express a causal relationship (and this aspect of the meaning is, naturally, a simple entailment), but only *min* has the added aspect that the Cause is dominant (and we have argued here that this aspect of the meaning is presuppositional).

2.6 Concluding remarks

In this chapter we have argued that the distinction between the causal uses of *bə* 'in' and *min* 'from' is one of dominance. When *min* is used, the argument is a dominant causer, whereas *bə* is unmarked, and gets to be used for non-dominant causers. Dominance in a causal model was defined as follows:

(2.20) A cause C of an effect E is represented as "dominant" if (a) C is exogenous (not dependent on other variables) and (b) E does not depend on any other exogenous variables besides C.

We believe our description of these prepositions to be an improvement over the traditional one, in which there was a significant overlap between the functions of the two (see section 2.2). This improvement was achieved by doing away with superimposed labels like 'Agent' and 'Instrument'.

Furthermore, we have shown how the fact that *min*, and not *bə*, marks dominance, can be derived from the difference in spatial meaning between the two prepositions. In a causal model, a dominant causer stands at maximum distance from the effect and at the origin of the dependency chain, since it is represented by an exogenous variable. It is not surprising that such a causer is marked by a preposition that also marks an Origin or Source in its spatial sense: *min*, as opposed to *bə*.[32] On the other hand, when used spatially, *bə* describes a physical relation with a smaller distance, and can as a result be used for causers closer to the effect in the causal model (e.g., an Instrument as opposed to an Agent; cf. a model like [2.22b]).

Though our main goal has been to describe and explain the distribution of causal *min* and *bə*, we finish with one example to show the exegetical value of our proposal:

(2.51) Exod. 15:23: וַיָּבֹאוּ מָרָתָה וְלֹא יָכְלוּ לִשְׁתֹּת מַיִם מִמָּרָה כִּי מָרִים הֵם עַל־כֵּן קָרָא־שְׁמָהּ מָרָה׃

way-y-ā̊ḇō'-ū	*mā̊rā̊-ṯā̊*	*wə=lō'*	*yā̊ḵəl-ū*	*li=štōṯ*
WAYQ-3M-come-PL	Marah-ALL	and=not	be_able\PFV-3PL	to=drink\INF

32 As mentioned before, I address the different spatial conceptualization of the causal chain in French (chapter 1) in chapter 6.

mayim	***mim***=*mā̆rā̆*	*kī*	*mā̆r-īm*	*hēm*	*ʿal kēn*
water(PL)	**from**=Marah	COMP	bitter-PL	they	therefore
qā̆rā̆'-Ø	*šəm-āh*	*mā̆rā̆*			
call\PFV-3M.SG	name-its	Marah			

'When they came to Marah, they could not drink the water *of* Marah because it was bitter; therefore it was named Marah.' (ESV)

In this verse, *lō' yā̆ḵəlū lištōṯ mayim mimmā̆rā̆* is typically translated as in the ESV, taking *mim-mā̆rā̆* with *mayim*: 'water of Marah'. It is also possible to read *min* as causal, if we take *mā̆rā̆* as an abstract noun: 'because of bitterness'.[33] The *kī*-clause can then not be read causally (because we already have a cause in *mim-mā̆rā̆*), but can be read as an exclamative clause instead (cf. section 4.7.3). The translation then becomes:

(2.51′) '..., they could not drink the water *from* bitterness — it was so bitter! — Therefore it was named Marah.'

While the traditional translation remains a good option as well, reading *mim-mā̆rā̆* as a causal prepositional phrase instead of as a locative phrase yields a more elaborate play on words, and the dominance marked by *min* resonates with the exclamative reading of *kī*. Our proposal on the difference between causal *min* and *bə* is therefore not only of theoretical importance, but should be taken into account by translators as well.

33 Cf. Gen. 26:35 (**mōrā̆*) and Prov. 14:10 (**mā̆rrā̆*), and for the use of the feminine for abstracts more generally see e.g. Joüon & Muraoka (2006: §134n).

Part II

Positioning information content

The interpretation of [+distal] in demonstratives and complementizers

3

Abstract This chapter argues that the [+distal] feature of demonstrative *that* is also present in complementizer *that*, and has not bleached away. In particular, we argue that complementizer *that* is referential: it refers to an element in the *Shared Discourse Space* (an extension of the Common Ground) that can be seen as distal. This allows us to explain (a) that direct speech patterns with [–distal] (*Sue said this*/#*that: "It is raining"*) while indirect speech patterns with [+distal] (*Sue said* **this*/*that it is raining*); (b) the use of *that* in exclamatives (*That bio industry is still allowed!*); and (c), that optional *that* is more frequently used when there is some sort of context between Speaker and Addressee. This last phenomenon has parallels in Romance complementizers derived from Latin *quod*, which can likewise be seen as [+distal]. We propose that [+distal] is a marker of *Addressee involvement*, which can account for all these phenomena, and can be extended to demonstrative uses of *that* as well. In exophoric contexts, [+distal] additionally marks *actual distance*. The interpretation of Addressee involvement and actual distance depends on context; we propose that it is derived from the interaction between the syntactic DP/CP domain and the pragmatic exophoric/endophoric distinction.

This chapter was originally published as: Camil Staps & Johan Rooryck. 2023. The interpretation of [+distal] in demonstratives and complementizers. *Linguistics* 61(5). 1195–1231. DOI: 10.1515/ling-2022-0178. Camil Staps and Johan Rooryck performed the analysis. Camil Staps wrote the article, in regular discussion with and with contributions by Johan Rooryck.

We are grateful to Ellen van Wolde and Joost Zwarts for useful feedback on this chapter. An earlier version of this chapter was presented in the SLE 2022 workshop on complementation and relativization, organized by Lena Baunaz, Tabea Ihsane, and Tania Paciaroni, and we thank the audience there for their comments. Finally, we are grateful for the useful feedback of two anonymous reviewers. Any remaining mistakes are, of course, our own.

3.1 Introduction

Indirect speech reports are commonly formed by a verb of saying and a finite complementizer introducing the sentential complement. In English, as in many other languages, this complementizer developed from a distal demonstrative, and it is indeed not possible to use a form with a proximal feature in this position. This suggests a close association between indirect speech reports and [+distal]:

(3.1) *Sue said* (**this*/*that*) *it is raining.* (cf. Rooryck 2019: 257)

The opposite is the case with direct speech reports. Direct speech is usually introduced only by a pause, or by quotation marks in writing (*Sue said: "It is raining"*). However, the speech report can be referred to with a cataphoric pronoun in the main clause.[1] This pronoun is then necessarily [−distal]:[2]

(3.2) *Sue said* (*this*/#*that*)*: "It is raining."* (cf. Rooryck 2019: 257)

How can we explain the relationship between direct/indirect speech and [−/+distal], respectively? Of course *that* in (3.1) and *this* in (3.2) have a different syntactic category, but that is irrelevant to our question since we are comparing the value of the [±distal] feature that *this* and *that* have in common. Simply claiming that the complementizer *that* is semantically bleached and entirely lacks a [±distal] feature is not sufficient; this simply shifts the question to the history of the form: why did *that*, and not *this*, develop into a finite complementizer (cf. Kayne 2014: 189)? Instead, we present a new, unified analysis which predicts a broad range of ways in which the proximal/distal distinction is recycled in both demonstrative and complementation environments, explaining the contrast in (3.1–3.2) as well as many other data adduced below.

The standard view on complementizers like *that* in (3.1) is that they fulfill a primarily syntactic function and are largely void in terms of semantics

1 Whether direct speech reports are subordinated or paratactic structures is inconsequential to our argument.

2 As an anonymous reviewer points out, *that* is possible in (3.2) if used anaphorically rather than cataphorically, e.g. *But he has said that: "Am I supposed to dislike them?"*. In such cases *that* is not coindexed with the speech report as *this* can be; instead, *that* refers to something in the previous context between Speaker and Addressee. We leave these cases out of consideration here.

and pragmatics, apart from carrying a feature indicating that they introduce a tensed rather than an untensed complement clause (Lasnik & Saito 1991: 324; Rizzi 1997: 312; implicitly in Rosenbaum 1965 and various grammars, e.g. Huddleston & Pullum 2002: 947–1030; and see discussion in Roberts & Roussou 2003: 111–116). In other words, mainstream theories of complementation do not attribute any synchronic value to the original distal semantics of the complementizer *that*, nor do they ascribe any other interpretively relevant information to it. However, a number of studies have indicated that these complementizers do carry additional interpretive information (e.g. Storms 1966; Bolinger 1972; Yaguchi 2001; Dor 2005). So far, such studies have mostly been restricted to rather specific contexts. Below, we will first draw attention to a number of recurring interpretive properties of complementizers and sketch the outline for a unified account. In this way, our account of the difference between *this* and *that* in (3.1–3.2) will also allow us to explain the difference between zero and *that* in English object clauses (*I thought* (*that*) *you might need some help*), the use of overt complementizers in exclamatives (*That bio industry is still allowed!*), and evidential interpretations of root complementizer constructions in Romance (to be exemplified below). We then show how these recurring properties can be explained as the interpretive recycling of a [±distal] feature. This allows for a general analysis of *this* and *that* covering both demonstrative and complementizer functions.

Concretely, we will argue that the proximal/distal distinction inherent in demonstratives can be recycled in two different ways, which we call *actual distance* and *Addressee involvement*.[3] The interpretation of these categories differs depending on the context. The picture that we are working towards is as in table 3.1. On the left we have demonstratives, which reference entities in the DP domain. These can be exophoric or endophoric (Diessel

3 For the term "recycling", see Rooryck (2019: 244), building on Biberauer (2017). What we mean by this is that markers of a certain category (here, proximal/distal) are repurposed to mark features of a different category (here, actual distance and Addressee involvement). This may be the first step in a grammaticalization process, in which the original deictic meaning has not been lost (yet). This perspective on *that* is thus quite different from the traditional view, which takes demonstratives and complementizers as *de facto* homonyms, at least synchronically (e.g. Diessel 1999: 123–125). It yields a more economical, polysemous view of demonstratives as exercising chameleon-like, distinct but strongly related functions, that vary according to the syntactic and pragmatic context in which they are used.

1999: 93–100). Exophoric demonstratives refer to something in the speech situation (*this*/*that book*), while anaphoric demonstratives refer indirectly through a linguistic antecedent in the surrounding discourse ([*Sales have been going up*]$_i$. [*This trend*]$_i$ ...). The complementizer *that* plays a role in the CP domain. We see it as referring to information content as opposed to entities in the speech situation. This reference can still be exophoric (when it refers to a concrete utterance, e.g. *Sue said that it is raining*) or anaphoric (when it refers indirectly through the Speaker's model of the discourse state, as in *I thought that you might need some help*; Bolinger 1972: 58).

	Entities (DP)	**Information content (CP)**
Exophoric	Exophoric demonstratives (section 3.4):	Direct and indirect speech (section 3.2):
	Actual distance in the concrete physical world	Actual distance in a multidimensional conceptual world, interpreted as similarity
	Addressee involvement: interpreted as psychological factors (psychological distance, joint attention, empathy, ...)	Addressee involvement: interpreted as evidentiality; proximity is private witness evidentiality
Anaphoric	Anaphoric demonstratives (section 3.5):	Presupposition (section 3.3):
	Addressee involvement: *that* used over *this* to interact and empathize with the Addressee	Addressee involvement: *that* used over zero to signal content in the Shared Discourse Space

Table 3.1 Deriving different kinds of reference from two binary properties.

We begin our discussion with reference to information content. In section 3.2, we use direct and indirect speech reports to introduce the notions of actual distance and Addressee involvement. Actual distance reflects the similarity between the speech report and the original utterance, whereas Addressee involvement is related to evidentiality. For the interpretation of Addressee involvement we introduce an extension of the notion of Common Ground, which we call *Shared Discourse Space*. The Shared Discourse Space, unlike the Common Ground, includes not only common commitments to propositions, but in broad terms all entities and information content that are jointly *tracked* by Speaker and Addressee as part of the dis-

course context (see section 3.2 for a more precise definition). Roughly, a direct speech report is more similar to the reported utterance than an indirect speech report (actual distance), and an indirect speech report places Speaker and Addressee on an equal footing with respect to the evidence for the reported utterance (Addressee involvement). We then move on to complementizers more generally in section 3.3, showing how Addressee involvement can explain alternations between overt and zero complementizers in a variety of environments (e.g., exclamative *that*, as in *That bio industry is still allowed!*, marks a presupposition that is shared with the Addressee). Sections 3.4 and 3.5 are dedicated to showing that the proximal/distal distinction is used in a similar way in demonstratives. Here, actual distance is simply physical distance to the object pointed at (*this*/*that book* being close to or far from the Speaker, respectively), and Addressee involvement concerns various psychological factors relevant to demonstrative choice. We show that [+distal] demonstratives, like [+distal] complementizers in the sentential domain, tend to be used more when the Addressee is more involved in the conversation. In section 3.6 we return to the matrix in table 3.1 to explain some gaps. In particular we answer the question why actual distance is not used with anaphoric reference and why *this* cannot be used as a complementizer. We also propose a definition of Addressee involvement that derives its interpretation in all four contexts in table 3.1. Finally, this section also discusses some related work and some final remarks.

3.2 Direct and indirect speech

As mentioned above, English allows direct speech complements to be introduced by the proximal demonstrative *this*, but not the distal demonstrative *that*. The latter has grammaticalized into a complementizer which can be used to introduce indirect speech, where *this* is not allowed:

(3.1) *Sue said* (**this*/*that*) *it is raining.* (cf. Rooryck 2019: 257)

(3.2) *Sue said* (*this*/#*that*)*: "It is raining."* (cf. Rooryck 2019: 257)

We argue that this pattern is not arbitrary, but is based on the recycling of the category of physical distance ([±distal]) in grammar. In the case of the distinction between direct and indirect speech, there are two target categories for the recycling process: *actual distance* and *Addressee involve-*

ment. Both provide a link between physical distance and the direct/indirect speech distinction.

We begin our discussion with actual distance, which is the most intuitive category in this context. Observe that direct and indirect speech reports differ in the degree to which the report is similar to the original utterance. Indirect speech reports do not need to be very similar to the original utterance; they only need to match their at-issue entailments, implicatures, and presuppositions (Brasoveanu & Farkas 2007). Thus (3.1) may for example be uttered after Sue has said something like *Why is it always raining when I want to go out?*; this original utterance matches in terms of entailments, implicatures, and presuppositions with the report in (3.1). However, it cannot be reported with the direct speech report in (3.2). For a speaker to faithfully utter (3.2), Sue's utterance must have been (almost) lexically identical to *It is raining*. It thus becomes clear that direct speech reports do not only have restrictions on the semantic and pragmatic content of the original utterance, but that they have additional constraints on its surface form. In this way, a direct speech report is more similar to the reported utterance than an indirect speech report. As a result, direct speech reports also lend themselves better to "personal" renderings of the original utterance, including the imitation of accents, pitch, accompanying gestures, etc. (Clark & Gerrig 1990). In this way direct speech again allows for greater similarity to the original utterance than indirect speech.

We think of this similarity in the following way. Both the original utterance and the speech report can be defined in terms of properties referring to their precise lexical form, propositional content, entailments, phonological information needed to represent accents, accompanying gestures, and possibly more features. This view of speech reports and utterances as multidimensional objects allows us to compare two of them and evaluate their similarity.[4] This is analogous to defining a point in the physical world with

4 This is similar to Paul Churchland's notion of a "state space" (also "similarity space"). Churchland proposes that "the brain represents various aspects of reality by a *position* in a suitable *state space*" (1986: 280; emphasis original). For example, a color can be defined as a point in a three-dimensional state space, where each dimension measures the degree to which one receptor type is activated. Colors can then be compared as similar or dissimilar by measuring the distance between them in this color space. Churchland proposes state spaces for different sensory systems. He also suggests that concepts can be represented in a state space for language use and propositional knowledge (1986: 299–306), which is what we attempt to do here.

x, y, and z coordinates and measuring the distance between two points. The difference is that utterances are represented in a multidimensional conceptual space rather than in a three-dimensional physical world. Nevertheless, this analogy shows that the similarity of a speech report to the original utterance can be seen as the recycling of the actual distance between the referent (the original utterance) and the deictic expression (*this* or *that* in the context of the speech report).[5] We will use the term *actual distance* to refer to the Euclidean distance both in the physical world and in the multidimensional conceptual space where the similarity of speech reports is assessed in terms of distance to the original. Note that it is also very common to talk about similarity in phonological or propositional form in terms of distance: *You think that's what he talks like? That doesn't even come* ***close!*** or *You couldn't be* ***further from*** *the truth.*

The second way in which the proximal/distal distinction is recycled is as Addressee involvement — and this category can be generalized to all other contexts that we will discuss. Addressee involvement is an interpretation of the "distance" between the referent (Sue's utterance) and the Speaker (of [3.1]–[3.2]). A direct speech report as in (3.2) is "close" to the Speaker, because its use suggests that the Speaker has direct, reliable knowledge of Sue's utterance. By the Speaker's uttering of (3.2), the Addressee also receives evidence for Sue's utterance, but it is only indirect evidence. The proximity expressed by *this* positions Sue's utterance close to the Speaker, and reflects that the Speaker has more direct evidence than the Addressee for Sue's utterance. The Addressee is much less involved. On the other hand, an indirect speech report as in (3.1) does not imply that the Speaker has direct evidence for the utterance. Speaker and Addressee can then share the indirect evidence: the evidence is in the Shared Discourse Space. Distal *that* positions the complement clause close to the Addressee because the Speaker and the Addressee have the same amount of evidence for the information in that clause, and the Addressee is more involved. Closeness to the Addressee is represented as distance from the Speaker, so a [+distal] element is used.

This view entails, perhaps counter-intuitively, that the Shared Discourse Space is distal for the Speaker. We see the Shared Discourse Space not as a

5 Throughout, we use the term "referent" for the thing to which the deictic expression refers (cf. Maes et al. 2022). This is different from "antecedent", since the referent is not normally a linguistic element but an entity in the speech situation (the physical book with *that book there*) or an utterance or proposition (as with speech reports).

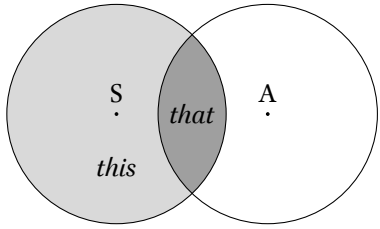


Figure 3.1 The Personal Discourse Spaces of the Speaker and the Addressee, with the Shared Discourse Space as their intersection.

region encompassing Speaker and Addressee, but as the intersection of their *Personal Discourse Spaces*: the collections of information content tracked by each of the interlocutors individually (including propositions, utterances, questions, ...). This is illustrated in figure 3.1. The Shared Discourse Space is therefore not proximal for the Speaker, but the proximal/distal distinction is used to distinguish between the information content private to the Speaker (proximal, light gray in figure 3.1) and the information content shared with the Addressee (distal, dark gray in figure 3.1). Therefore, although the speech report is positioned either close to or far from the Speaker, this is actually used to mark its absence or presence in the Addressee's Personal Discourse Space, respectively. For this reason we speak of *Addressee involvement* with a focus on the Addressee rather than the Speaker. In the case of speech reports, this Addressee involvement receives an evidential interpretation, with proximity/distance to the Speaker being recycled for direct/indirect evidentiality. As we shall see below, Addressee involvement receives a different interpretation in other contexts.

It is useful to briefly compare our model of the Shared Discourse Space to approaches to the Common Ground (e.g. Lewis 1969; Stalnaker 1978; Clark 1996; Farkas & Bruce 2010), which is similar to the Shared Discourse Space (but remains useful for the explanation of other phenomena). First, the notion of Shared Discourse Space is broader than that of Common Ground. For Farkas & Bruce (2010) the Common Ground is the intersection of the interlocutors' commitment sets, and the commitment set of an interlocutor consists of the propositions she has publicly committed to. Because the commitment set is defined in terms of public commitments and not knowledge or belief, conversation participants know the contents of each other's commitment sets. But the focus on commitment to propositions makes this model too constrained for our purposes. For our analysis of demonstratives

in sections 3.4 and 3.5 it will be necessary to also include referents of demonstratives in the Shared Discourse Space. Furthermore, the sensitivity of complementizers to previous questions or utterances explored in sections 3.3.2 and 3.3.3 also requires a notion broader than commitment to propositions. For this reason, the Shared Discourse Space does not (only) contain information about commitment to propositions, but more generally about all the entities and information content that are tracked by the interlocutors as part of the discourse context. The Personal Discourse Space of an interlocutor consists of the entities and information content that are tracked by her — and we take *tracking x* to be general enough to include believing *x*, believing that ¬*x*, being interested in whether *x* is the case, pondering the requirements or corollaries of *x*, having any kind of emotional attitude towards *x*, etc. It therefore includes, but is not limited to, the Common Ground. It is also not limited to propositions: it may contain *x* if the interlocutor tracks that *x* has (or has not) been uttered, that *x* did (did not, might, should, etc.) occur, or how *x* can be identified. The Personal Discourse Space contains the entities and information content to which an interlocutor is, in a broad sense, attentive. The Shared Discourse Space, then, consists of the entities and information content that are tracked by the Speaker and that are assumed (by the Speaker) to be tracked by the Addressee.[6]

We approach the Shared Discourse Space explicitly from the point of view of the Speaker: the Shared Discourse Space is the intersection of that which the Speaker considers her Personal Discourse Space and that which she assumes is the Addressee's Personal Discourse Space. Others already recognized the need for the perspective of the Speaker in the analysis of Common Ground. For instance, Clark (1996: 96) notes that only an omniscient being can say "It is common ground for the two of them that [...]", and conversation participants can only say "I believe that it is common ground for us that [...]". Clark recognizes that there may be situations where the interlocutors have different ideas of what the Common Ground contains. In such situations, the language used by the interlocutors is determined by the assumptions they make about the Common Ground — not by what an om-

6 Note that this derives the fact that there is no extra-distal demonstrative to refer to something in the Personal Discourse Space of the Addressee but not in that of the Speaker: as soon as something in this region is referred to by the Speaker, it becomes part of her Personal Discourse Space as well, because it becomes tracked. There is no way for the Speaker to talk about something without tracking it herself.

niscient being would theoretically know that the Common Ground consists of. The same applies by extension to the Shared Discourse Space. Stalnaker (1978: 321) seems to recognize the same thing when he writes that "presuppositions are what is *taken by the speaker* to be the common ground" (emphasis ours), but later defines Common Ground without taking into account the perspective of the conversation participants: "the common beliefs of the parties to a conversation are the beliefs they share, and that they recognize they share" (Stalnaker 2002: 704). This type of Common Ground only exists as a theoretical construct, it is inaccessible to the interlocutors and therefore cannot influence the way they speak. For this reason, we explicitly take the Speaker's perspective on the Shared Discourse Space.

The treatment of speech reports and demonstratives proposed in this section has many precursors in the literature. For instance, Clark & Gerrig (1990: 792–793) observed that the Speaker of a direct speech report takes responsibility for the correct rendering of an utterance, while the Speaker of an indirect speech report takes responsibility for the interpretation of an utterance. Wierzbicka (1988: 132–135) has an analysis of indirect speech *that* which is similar to ours, although she compares it to direct speech introduced with a pause rather than proximal *this*. She argues that direct speech reports "sound like reports of utterances expressing emotion, rather than 'objective' judgement", while indirect speech reports "imply that the speaker was trying to assess the reality, not merely to express his emotion" (1988: 133). For instance, utterances that are high in emotive attitude, like *You idiot!*, can hardly be reported with indirect speech (?*He said that she was an idiot*; preferred would be: *He called her an idiot*). This can be seen as a reluctance to refer to the meaning of emotive utterances as opposed to the utterance itself. This reluctance would be understandable: if the original Speaker made an emotive utterance, she may not be held fully responsible for its propositional content because the utterance may be made in the heat of the moment. However, neither Clark & Gerrig (1990) nor Wierzbicka (1988) related these observations to the proximal/distal distinction that remains present in complementizers. Rooryck (2019: 256–257) does discuss speech reports with reference to the proximal/distal distinction, but considered the Common Ground to be proximal to the Speaker.[7]

7 In Rooryck (2019: 256) it was suggested that proximal *this* places the content of a direct speech report in the Common Ground, because the Common Ground is proximal to Speaker and Addressee. By contrast, we take *that* to involve reference to the Shared

3.3 Presupposition effects

Having shown in the previous section how the [±distal] feature is recycled to mark actual distance (interpreted as similarity) and Addressee involvement in the context of speech reports, we now turn to cases where overt complementizers contrast with zero complementizers.[8] In these cases there is no difference in terms of actual distance, but the notion of Addressee involvement does generalize.[9] Our position will be that overt complementizers which are historically based on non-proximal elements markedly involve the Addressee. In particular, we analyze the examples below using the notion of Shared Discourse Space. When information content is in the Shared Discourse Space, it is shared with the Addressee, and therefore "far" from the Speaker; when information content is not in the Shared Discourse Space but is tracked by the Speaker alone, it is instead "close" to the Speaker. The proximal/distal distinction is thus recycled to indicate the absence/presence of content in the Shared Discourse Space.

3.3.1 Exclamatives

We first look at main clauses with overt complementizers, which in many languages can get an exclamative reading:[10]

(3.3) a. *That bio industry is still allowed!*
b. *That he should have left without asking me!*
(Quirk et al. 1985: 841 in citation by Zevakhina 2013: 167)

Discourse Space while *this* refers to the Personal Discourse Space, i.e., to information content tracked by the Speaker but not shared by the Addressee. Therefore, for reasons outlined above, direct speech reports are not placed in the Shared Discourse Space but remain personal to the Speaker, while indirect speech reports are shared with the Addressee.

8 Depending on one's syntactic framework, these cases could also be analyzed as contrasts between an overt complementizer and the lack of a complementizer. This does not affect the argument: in the end it is the presence (or absence) of the [+distal] feature that matters. We will use "zero complementizer" for simplicity, without making any theoretical assumptions.

9 We return to the question why actual distance is not relevant here in the conclusion.

10 However, constructions in Romance of the type *Que cette histoire est obscure!* 'How dark this story is!' (French) should be kept separate, because they always refer to a degree rather than a fact (Trotzke & Villalba 2021). We are grateful to Maria Bardají i Farré for suggesting this reference.

(3.4) Swedish (Delsing 2010: 17 in citation by Zevakhina 2013: 167)

Att du hann med tåg-et!
COMP you reach.PST with train-the

'(It is surprising,) that you caught the train!'

(3.5) Polish (based on Storms 1966: 261[11])

Że też potrafiłeś coś takiego zrobić!
COMP also can.PST.2M.SG something this do.INF

'That you could do something like this!'

(3.6) Biblical Hebrew
Gen. 18:20:[12] זַעֲקַת סְדֹם וַעֲמֹרָה כִּי־רַבָּה וְחַטָּאתָם כִּי כָבְדָה מְאֹד׃

za'aqa-ṯ sədōm wa='ămōrå̄ kī råbb-å̄
outcry(F)-of Sodom and=Gomorrah COMP be_great\PFV-3F.SG

wə=ḥaṭṭå̄ṯ-å̄m kī ḵå̄bəḏ-å̄ mə'ōḏ
and=sin(F)-theirs COMP be_heavy\PFV-3F.SG very

'The outcry of (/against) Sodom and Gomorrah, how great it is! And their sin, how very grievous!'

In these examples, the exclamative is only distinguished from a regular declarative sentence by the addition of the complementizer and a different intonation pattern. The intonation pattern alone is not enough for the exclamative interpretation. For instance, a sentence like *Bio industry is still allowed!*, with the same intonation pattern as the exclamative, still differs from an actual exclamative like (3.3a) in that it can be used to attempt to convince the Addressee of its propositional content. By contrast, the sentence in (3.3a) does not make an attempt at informing or convincing the Addressee of its propositional content, but actually presupposes it to be a shared presupposition in the Common Ground, and hence in the Shared Discourse Space. The use of the complementizer is therefore crucial for the interpretation as an exclamative. Zanuttini & Portner (2003) already showed that exclamatives are factive.[13] On this view, exclamatives make ref-

11 We are grateful to Justyna Visscher-Jablonska for providing a Modern Polish version and glosses.

12 On this example see further section 4.7.3.

13 In terms of Ross's (1970) performative hypothesis, factivity of exclamatives would be explained through the deletion of a factive performative (*I am surprised that ...!* > *that ...!*); see also Evans (2007) for diachronic considerations. In neo-performative treatments the performative structure is not deleted but part of the functional domain above

erence to a proposition that is already presupposed in the Shared Discourse Space and relate a certain Speaker stance (surprise, anger, etc.) to it.[14] We propose that the [+distal] complementizer in these exclamatives anaphorically refers to the presupposed proposition.[15]

In these cases, the referent (the presupposition in the Shared Discourse Space) is always "far" from the Speaker: we do not find exclamatives with a complementizer or other grammatical marker specified for [–distal].[16] The notion of Addressee involvement makes it easy to see why: if, following Zanuttini & Portner (2003), exclamatives require presupposition, they must refer to the Common Ground, and hence to the Shared Discourse Space. An exclamative cannot at the same moment introduce new, Speaker-personal information content into the discourse. As a result, the information con-

---

the CP (e.g. Speas & Tenny 2003). Alternatively, *that* could be seen as an underspecified element, with its factive meaning deriving from the merge site (cf. Kocher 2022 on Ibero-Romance *que*). The exact derivation of exclamatives is not relevant here; what is important is primarily the fact that exclamatives are factive.

14 In some cases, the proposition is strictly speaking not presupposed but can be easily accommodated by all interlocutors. We see such cases as involving an imposition on the Common Ground, and by extension on the Shared Discourse Space, through referencing a proposition: by referencing the proposition, the Speaker pretends that it is already in the Common Ground, thereby imposing an update to the Common Ground. Also see our discussion of Kocher (2022) in section 3.3.3 below, especially footnote 26.

15 Note that other syntactic strategies of exclamatives also often contain anaphoric elements:

(i) Mandarin Chinese (Visan 2000: 9 in citation by Zevakhina 2013: 169)

Zhège háizi!
this child

'What a child!'

(ii) *It's so hot!* (Michaelis 2001: 1040 in citation by Zevakhina 2013: 166)

(iii) Russian (Zevakhina 2013: 166)

Miša takoj bol'šoj!
Miša such.NOM big.NOM

'Miša is so big!'

Although these anaphoric elements do not refer to a presupposed proposition, they still establish Shared Discourse Space between Speaker and Addressee.

16 See example (i) in footnote 15 for a case where a proximal element can be used in an exclamative. But note that this is an exophoric demonstrative and does not head the exclamative clause.

tent must be close to the Addressee, and therefore a distal element must be used.[17]

3.3.2 The *that*/zero alternation in English object clauses

We can also use Addressee involvement to explain the alternation between overt and zero complementizers in English object clauses. Consider (3.7):

(3.7) a. *I thought you might need some help.* (Bolinger 1972: 58)

b. *I thought that you might need some help.* (Bolinger 1972: 58)

A common view is that the complementizer *that* in (3.7b) is "optional", i.e., that its use is determined by style or register and that it does not have an interpretive value. However, the literature discusses many factors that can play a role in the choice between *that* and a zero complementizer. Two in particular suggest that we are actually dealing with an interpretively meaningful alternation and not with an entirely optional functional element.[18]

17 Ellen Brandner (p.c., August 26, 2022) notes that in some German exclamatives the distal demonstrative *der* is preferred over the personal pronoun *er* 'he', as in *der*/*#er und lesen!* 'he and reading!' (i.e., he will definitely not read; the idea is preposterous). The preference for the distal demonstrative cannot be explained by emotional distancing from the subject, as the same effect appears with predicates with a negative connotation: *der*/*#er und Plagiat begehen!* 'he and committing plagiarism!' (i.e., he will definitely not commit plagiarism, which is a meliorative statement and would not require emotional distancing). The affinity of exclamatives with distal elements may thus extend beyond the complementizer.

18 We are not concerned here with cases where *that* is used to avoid ambiguity or otherwise make parsing the sentence easier (e.g., Bolinger 1972: 18–42; Elsness 1984). Beal (1988: 60) and Rissanen (1991) observed that *that* is more often omitted in constructions that frequently take complement clauses, because the pattern is less unexpected and does not need to be marked by *that*. We take this to indicate that *that* is inserted in infrequent collocations to clarify the sentence structure (cf. also Kajzer-Wietrzny 2018). Although these factors are not relevant to us here, one should be aware of their existence because they can interfere with minimal pairs. We also set aside here style and register (Elsness 1984; Rissanen 1991), as well as the suggestion found in Hooper & Thompson (1973), Thompson & Mulac (1991), Diessel & Tomasello (2001), and Thompson (2002) that certain combinations of first and second person subjects and verbs like *think* and *guess* can be reanalyzed as markers of epistemic modality so that the distinction between main and complement clause erodes and *that* is less likely (see also Kaltenböck 2009 and Dehé & Wichmann 2010, who seek to predict when clause-initial constructions like *I think* (*that*) are such epistemic markers, as opposed to matrix clauses).

Firstly, Bolinger (1972: 58) already noticed that the sentence with *that* in (3.7b) suggests some context between Speaker and Addressee. This context may be extralinguistic, as in the scenario he sketches:[19] "Suppose you observe a stranger struggling to mount a tire. Feeling charitable you go over to him and say [3.7a]. Under these circumstances, [3.7b] would be inappropriate. But if the other person looks at you as if wondering why you came over, you might explain by saying [3.7b]" (Bolinger 1972: 58, example numbers adapted).

In the words of Bolinger (1972: 56), the complementizer still "reflects the demonstrative character of *that*" in that it refers to this shared context. After all, this use of *that* appears to be quite similar to the discourse deictic function of demonstratives (e.g. *That's a lie*; Diessel 1999: 101). Both refer to some utterance, even though the utterance is only implied in (3.7b) (i.e., we assume there to be an implicit utterance along the lines of *Why did you come over?*). The situation is then quite similar to that of exclamatives: the use of an overt complementizer signals content in the Shared Discourse Space. Again, then, the referent (the presupposed utterance) is analyzed as "far" from the Speaker, triggering a distal element, because it is in the Shared Discourse Space, close to the Addressee. In (3.7a), no anaphoric element is present because the idea that the Addressee might need help has not yet been introduced, and is therefore not in the Shared Discourse Space. The presence or absence of the complementizer thus marks the presence or absence of shared context in the Shared Discourse Space.[20]

The other relevant factor conditioning the choice between *that* and zero is that of subjectivity (Storms 1966: 262–265). Storms argues that sentences incorporating a *that*-clause are "less personal, less familiar, less warm, less

19 Bolinger (1972) gives many more examples. Some native speakers we consulted did not share Bolinger's intuition expressed here. This may be due to the fact that there are many different factors that play a role in the choice between *that* and zero. The relative weight of these factors could vary between variants of English, speakers, and contexts. More work is needed to establish the extent to which Bolinger's intuitions are still relevant today. For the present study, it suffices to say that shared context between Speaker and Addressee played a role in at least one variety of English at one point in time.

20 According to Auer (1998), in citation by Weinert (2012), something similar is the case with German *dass* 'that': "unintroduced main clauses are relatively assertional (they tend to contain fore-grounded and new information) whereas introduced complement clauses are relatively presuppositional (they tend to contain back-grounded and known information)" (Weinert 2012: 243).

friendly, less emotive" than their counterparts with zero complementizers (Storms 1966: 262). He gives examples from a witness interrogation in court, where sentences without *that* are used "to put the witness at her ease and at the same time to set an unsuspected trap" (Storms 1966: 263). Later, when it is important that objective facts are established, questions with *that* are used (Storms 1966: 264). Similar ideas appear in Wierzbicka (1988: 132–140), who relates *that*-clauses (as opposed to other complementation types) to knowledge. We believe that this subjectivity derives from the placement of the complement in or outside of the Shared Discourse Space. The lawyer cited by Storms (1966) uses *that* for propositions that are not yet in the Common Ground, but by using *that* he implicitly proposes to update the Common Ground to include them.[21]

Previously, Kaltenböck (2006) already proposed to use the abstract notion of distance to explain the difference between *that* and zero in extraposed *that*-clauses (as in *It is obvious* (*that*) *she did it*), suggesting that the analysis could be generalized to object clauses as well (Kaltenböck 2006: 389 n. 20). In his view, the abstract notion of distance is interpreted as one or more of (a) illocutionary distance (asserting the complement with zero vs. disposing the matrix for illocutionary force with *that*), (b) temporal/anaphoric distance (using *that* for complement clauses whose content has already been talked about vs. zero for new information), and (c) emotional distance (à la Storms 1966). However, Kaltenböck (2006) does not provide a principled reason why old information should be distal (in terms of temporal/anaphoric distance); we might as well argue that discourse-old information is proximal, because what is close to us is better known than what is far. The notion of Addressee involvement provides an explanation: discourse-old information is distal because it is in the Shared Discourse Space, known and shared by the Addressee. Addressee involvement is also needed to explain the lack of a proximal complementizer *this*, which Kaltenböck's (2006) analysis does not seem to predict. We return to this issue in section 3.3.4. We discuss more related work in section 3.6.2.

To finish our discussion of optional *that* in object clauses we briefly discuss the fact that *that* is required when the object clause is topicalized:

21 The use of *that* in "less friendly" contexts could also be related to the formal register with which *that* is associated. However, conversely it may also be the case that *that* is associated with formal language precisely because of this interaction with "subjectivity".

(3.8) a. *I always believed* (*that*) *the jury was bribed.*

b. *(*That*) *the jury was bribed, I always believed.*[22]

As discussed by Rizzi in Kratzer et al. (2020), there have been different syntactic accounts of this phenomenon. In this interview, Rizzi proposes an account that adopts an idea from Pesetsky (1995). Rizzi assumes that the lack of a complementizer indicates incorporation or cliticization of that complementizer into the selecting verb. Since the C head of the complement clause has already moved in (3.8b), the complementizerless clause cannot in turn move to a higher position: the complement clause is frozen in place. This would explain the pattern in (3.8). However, this account needs to introduce the otherwise uncorroborated assumption that complementizers incorporate into the selecting verb in English. By contrast, the analysis based on Shared Discourse Space that we present here suggests an explanation that derives from a wider generalization: *that* is required in topicalized object clauses because topics are necessarily discourse-old, and hence in the Shared Discourse Space. The complementizer *that* is then required to indicate the shared status of the topicalized clause.

3.3.3 Overt root complementizers in Romance

Finally, presupposition and the Shared Discourse Space also play a role in constructions with a root complementizer found in several Romance languages. It may not be immediately obvious that the complementizers discussed here contain a non-proximal deictic element, but we return to this issue below. One type of root complementizer construction that we are in-

22 *That*-deletion is actually obligatory in this example if *I always believed* is taken as an evidential modifier (cf. *The jury was bribed, I think*). However, in such a case *that* is not permissible because *The jury was bribed* is the main clause (cf. Bolinger 1972: 15–16, 62; Hooper & Thompson 1973; Thompson & Mulac 1991). The two can be distinguished by the fact that the sentence with an object clause does not make an attempt to update the Common Ground: (3.8a) is acceptable in contexts where the Addressee does not need to accept that the jury was bribed. For instance, *whatever* in *I always believed the jury was bribed, but whatever* indicates that the Speaker does not care about the Addressee's commitment to the proposition that the jury was bribed, whereas *whatever* in *The jury was bribed, I always believed, but whatever* indicates that the Speaker does not care about the fact that the jury was bribed. We are concerned here with the sentence with topicalization, which does make an attempt to update the Common Ground and in which *that* is required.

terested in here has a sentence-initial adverb followed by an overt complementizer:[23]

(3.9) Spanish (Etxepare 1997: 98–99 in citation by Hernanz 2007: 165–166)

Evidentemente (*que*) *Julia está muy enfadada.*
obviously COMP Julia is very angry

'Obviously, Julia is very angry.'

According to Etxepare (1997: 99 in citation by Hernanz 2007: 166), the felicity of Spanish *que* in sentences like (3.9) is conditioned by the occurrence of a clear "linguistic antecedent" in the preceding discourse (cf. Etxepare 2010: 613). Thus, (3.9) is only felicitous after another Speaker has uttered a sentence like *¿Se ha enfadado Julia?* 'Did Julia get angry?'. We might analyze this linguistic antecedent as establishing a Question Under Discussion (QUD) in the Shared Discourse Space.[24] The existence of this QUD then licenses the use of *que*.

We can conclude that, as in the cases discussed above, the use of an overt complementizer is licensed by the existence of an element in the Shared Discourse Space (namely, the QUD). Because we want to focus here on the properties of the Shared Discourse Space, we do not go into details about a possible formal representation of the QUD.[25] What is relevant to us is only that the Shared Discourse Space (a) contains informational elements, (b) that these elements can be tracked and referred to by conversation participants, and (c) that some of these elements can be marked as being under discussion.

According to Kocher (2022: 75–82), the linguistic antecedent requirement is not as strict as assumed by Etxepare (1997, 2010), as *que* can also be used for future or hypothetical utterances:

23 Other, similar constructions are discussed in depth by Kocher (2022: 91–196). She shows that all these cases impose a commitment to a proposition on the part of the Addressee, which is a clear case of Addressee involvement and very similar to the cases discussed above. For reasons of space we discuss only one construction here.

24 This treatment is similar to that of Pérez & Verdecchia (2022) for "clausal doubling", which covers cases like: *Que leyó el libro, seguro que lo leyó* 'As for her reading the book, she read it for sure'. The first clause is seen as establishing the QUD; the second clause has *que* because it responds to this QUD.

25 See e.g. Ginzburg (1996), Roberts (2012 [1996]), Büring (2003), and Farkas & Bruce (2010).

(3.10) Catalan (Kocher 2022: 77)

Avisa el comissari. Que ja pot venir.
notify.2SG.IMP the inspector COMP already can.3SG.PRS come
'Notify the inspector. [reportative:] He can already come.'

In Kocher's (2022) analysis, *que* merges in a high position in the left periphery in cases like (3.10), where it simply indicates that the sentence is subordinate. The Addressee can then infer that a *verbum dicendi* is implicitly understood. This is in contrast to cases like (3.9), where *que* merges in a low position in the left periphery where it expresses that a commitment to the proposition is attributed to the Addressee; an attributive feature in the sense of Poschmann (2008). This attributive feature, which is a form of Addressee involvement, explains why B's response is felicitous in (3.11a) below but not (3.11b). In (3.11b), the attributive feature of *que* clashes with A's sentence in which the proposition is described as a false belief:

(3.11) Spanish (Kocher 2022: 175–176)

a. A: *Qué dicen los doctorandos al inicio de sus estudios?*
what say.3PL.PRS the PhD_students at_the beginning of their studies
'What do PhD students say at the beginning of their studies?'
B: *Que {seguro que / seguramente} acabarán su tesis a tiempo.*
that sure COMP surely finish.3PL.FUT their thesis on time
'That surely they will finish their thesis on time.'

b. A: *Cual es la falsa idea que tienen los doctorandos al inicio de sus estudios?*
what be.3SG.PRS the false idea that have.3PL.PRS the PhD_students at_the beginning of their studies
'What is the false belief that PhD students have at the beginning of their studies?'
B: *Que {#seguro que / seguramente} acabarán su tesis a tiempo.*
that sure COMP surely finish.3PL.FUT their thesis on time
'That surely (#*que*) they will finish their thesis on time.'

We are agnostic towards the exact syntactic derivation leading to the attested surface structures. What is important to us here is that in both cases distinguished by Kocher (2022), the use of *que* is conditioned by the existence of shared context between Speaker and Addressee. In cases like (3.9), the proposition itself is placed in the Shared Discourse Space; in cases like (3.10), there is the salient *verbum dicendi* that provides the shared context.[26]

We find similar constructions in other Romance languages:[27]

(3.12) Romanian (Cruschina & Remberger 2017: 89)

Sigur (*că*) *va* *veni.*
sure COMP will.3SG come

'Of course s/he's coming.'

In (3.12), Romanian *că* may only be used when the Addressee could have inferred the propositional content of the clause. Cruschina & Remberger (2017: 89) set up the following contexts. Suppose Ioana asks Alexandru if Ion will attend a conference next week. Ioana does not and cannot have this information, but Alexandru has spoken to Ion and knows that he is coming. Alexandru can then answer with *Sigur va veni.* However, suppose now that Alexandru does not have this information, but that both Ioana and Alexandru know that Ion is a big fan of the conference and would never miss it. In this context, Alexandru can answer with *Sigur că va veni.* The answer is then marked as an inference from information in the Common Ground between Ioana and Alexandru, rather than as private information of Alexandru. Again, we see that reference to the Shared Discourse Space (here in particular the Common Ground), and hence Addressee involvement, is marked by an overt complementizer.

26 A small difference between our account and that of Kocher (2022) is that she describes the pragmatics of attributive *que* as "imposing" a proposition on the Common Ground, whereas we talk about referencing a proposition in the Common Ground (and by extension in the Shared Discourse Space). The term "reference" could seem to suggest that the proposition must already be in the Common Ground. This is, of course, not the case: it is perfectly possible to introduce new information in *que*-clauses. We suggest that this information is introduced by referencing it. The speaker "imposes" it on the Common Ground by pretending that it is already there. We hold on to the term "reference" to highlight the parallels with demonstratives, discussed below. The difference with "imposition" is largely terminological.

27 We use an example from Romanian here; for examples from other languages see Cruschina & Remberger (2017). For more discussion on Romanian, see Hill (2012).

A related phenomenon is found in Neapolitan. The following contrast is discussed by Sornicola (1996: 334–336) and Ledgeway (2011: 286–289):

(3.13) Neapolitan (Ledgeway 2011: 286[28])

a. *Chillo*$_i$ *s'è* *astutato* [*'o* *riscaldamento*]$_i$
that.M self=is turned_off the.M.SG heating.SG
'The heating has gone off.'

b. *Chello*$_i$ *s'è* *astutato* [*'o* *riscaldamento*]$_j$
that.N self=is turned_off the.M.SG heating.SG
'(The fact is/Because) the heating has gone off.'

On Ledgeway's double subject analysis, *chello*/*chillo* is not a complementizer but a demonstrative, but it is still similar to the cases discussed above. In (3.13a), *chillo* is coreferential with the second subject ('It has gone off, the heating'). In (3.13b), neuter *chello* cannot be coreferential with masculine *'o riscaldamento*. The demonstrative must therefore refer to something else. It has "a distinctly explicative or adversative value, only proving felicitous in contexts that contain an implicit or explicit presupposition" (Ledgeway 2011: 287). We suggest that in (3.13b) the demonstrative refers to this associated presupposition, as is the case with the complementizers in Spanish, Catalan, and Romanian.

We have largely left French aside in the discussion above. The reason for this is that French *que* is nearly obligatory in all environments. There appear to be some varieties that do allow *que* to be dropped in some contexts,[29] but these cases have not been described in sufficient detail yet to be included in our discussion here.

We should pause here for a moment to reflect on the origin of these complementizers. Above, we argued that [+distal] complementizer *that* marks the use of Shared Discourse Space because the Shared Discourse Space includes the Addressee and is therefore seen as "far" from the Speaker. However, the Romance complementizers discussed in this subsection are not demonstrative synchronically, so how does a [±distal] feature fit in? Note

28 The indices in (3.13b) have been corrected from the source after consultation with Adam Ledgeway (p.c., June 16, 2022).

29 Tabea Ihsane (p.c., August 25, 2022) kindly shared an observation that in some modern varieties of French it does seem to be possible to drop *que* under certain circumstances. We also thank Alina McLellan for discussing the situation in Réunion Creole, where *ke* (< Fr. *que*) is optional in many contexts (cf. Corne 1995).

that these complementizers derive from Latin *quod*, which is composed of an interrogative element (*qu-*) and the originally neuter medial demonstrative *id*. Given the latter component, these complementizers do diachronically derive from a non-proximal demonstrative element. There is also reason to believe that the interrogative element *qu-* is incompatible with proximity. For instance, consider that English has *what* from *that* and *where* from *there*, but not **whis* from *this* or **where* with an /ɪ/-vowel from *here*. Rooryck (2003: 11–12) suggests that this is because something that is proximate to the Speaker is necessarily known to them. In this way the interrogative element *qu-* could also be seen as a [+distal] component, thus involving the Addressee.

Our analysis of these complementizers is very similar to that of exclamatives. In (3.12), the sentence without *că* has an "objective" interpretation ('It is certain that s/he's coming'), whereas *că* triggers a "subjective, speaker-oriented" interpretation ('Of course s/he's coming'), where the propositional content is inferred (Cruschina & Remberger 2017: 88–89). This Speaker-oriented interpretation uses *că* to refer to a presupposition, just as exclamative complementizers refer to the proposition presupposed by their complement. This is entirely in line with Gutiérrez-Rexach (2001: 184–186), who calls these sentences in Romance "evidential exclamatives" and analyzes them as in (3.14a):

(3.14) Spanish evidential exclamatives (Gutiérrez-Rexach 2001: 184–185)

a. $[_{\text{Force}}\ \text{Adv/Adj}_{[+\text{evidential}]}\ [_{\text{Focus}}\ [+\text{f}]\ [_{\text{Topic}}\ \text{C}\ \ldots]]]$

b. *¡Claro que te lo voy a dar!*
clear that to_you it go to give
'Of course I will give it to you!'

According to Gutíerrez-Rexach, the evidential adverb *claro* 'clearly, of course' requires that its complement references a Question Under Discussion (QUD). For example, (3.14b) may be uttered if the Speaker has borrowed something from the Addressee, and the Addressee has expressed doubts about getting it back. The QUD is topicalized by the complementizer *que*. Because the complementizer is demonstrative, the QUD does not need to be spelled out, but the complementizer does need to be overt. The complementizer effectively points to the QUD in the Shared Discourse Space. Note, however, that the fact that it points to the QUD (and not any other element of the Shared Discourse Space) appears to be a language-particular

constraint: it applies in Spanish, Catalan, and Neapolitan, but not in Romanian, where *că* does not refer to a QUD but to any evidential basis for the claim made in the complement clause. What is at issue for us here is the generalization that the complementizer points to an element of the Shared Discourse Space.

Example (3.14b) illustrates the division of labor between the sentence-initial adverb and the complementizer, as well as the parallel with the exclamatives discussed in section 3.3.1. As with exclamatives, the function of the complementizer is to mark the existence of Shared Discourse Space between Speaker and Addressee. The sentence-initial adverb in the Romance root complementizer constructions only specifies the evidential interpretation.[30] The [+distal] element is therefore again used to signal Addressee involvement.

3.3.4 Presupposition effects: summary

To summarize the findings from this section: evidence from a variety of constructions (exclamatives, English "optional" *that*, and the root complementizer constructions in Romance) suggests that the alternation between an overt complementizer with a [+distal] feature and a zero complementizer is related to presupposition. We explain this by suggesting that the complementizer refers to information content in the Shared Discourse Space. Distal elements are used in these complementizers because the Shared Discourse Space includes the Addressee, who is "far" from the Speaker. Note that the theory correctly predicts that we do not find [–distal] elements in these environments, that is, that there is no complementizer derived from the demonstrative *this*. These would correspond to presuppositions that are not shared with the Addressee; a contradiction in terms, since presuppositions are necessarily assumed to be shared by all interlocutors. The only available alternation is with a zero complementizer, which marks the ab-

30 Note that Cruschina & Remberger's (2017) term "Speaker-oriented" for these evidential sentences refers to the fact that the Speaker makes an inference on the basis of the presupposed proposition. The proposition itself is presupposed, and therefore necessarily not Speaker-oriented, but shared between Speaker and Addressee. This leads to the odd situation that a distal element, which is typically used to trigger a more objective interpretation by placing something in the Shared Discourse Space, actually triggers a "Speaker-oriented" reading.

sence of a presupposition from both the Shared Discourse Space and the Personal Discourse Space.

Our analysis raises questions for the traditional account of the grammaticalization of *that* and cognate complementizers. The traditional view is that *that* became a complementizer as a result of reanalysis of a cataphoric demonstrative: *I say that: he comes* > *I say that he comes* (e.g. Diessel 1999: 123–125; Roberts & Roussou 2003: 113–120). As a cataphoric demonstrative, *that* introduces new information, which would be consistent with the use in *I say that: he comes*. But the shift to *I say that he comes* would be odd if the complementizer *that*, as in our analysis, refers to Shared Discourse Space (as opposed to introducing new information).[31] However, recent studies have suggested that the complementizer *that* instead developed from a correlative construction: *I say that, that he comes* (e.g. Axel-Tober 2017, and see Bate in preparation for a survey of finite complementizers in Indo-European). In such a construction, the first pronoun introduces new information but the second can be seen as referring to the Shared Discourse Space (as established by the first pronoun). This grammaticalization path therefore does not suffer from the same problem. Our analysis provides further support for this development.

Finally, although our focus here has been on complementizers derived from demonstrative pronouns, the phenomenon that finite complementizers are related to presupposition seems to be more general than that. For example, the Bulgarian relativizer *deto* (lit.: 'where the', i.e. 'the place where') is also used to express Speaker stance about presupposed propositions: *Sǎžaljavam, deto ne možax da dojda* 'I regret that I couldn't come' (Krapova 2010: 1240). It may therefore be that a finite complementizer does not need to be derived from a demonstrative pronoun, but that any deictic origin would suffice. We will not explore this further here.

3.4 Exophoric demonstratives

In sections 3.2 and 3.3 we examined the complementizer *that*. We compared this functional element to both the proximal cataphoric demonstrative *this* (for direct speech) and a zero complementizer (in main and object

31 Another issue with this diachronic account is that *I say that: he comes* is less natural than *I say this: he comes*, while *this* did not grammaticalize into a complementizer (Kayne 2014: 189).

clauses). Both sections were concerned with reference to information content, namely, the meaning of utterances (which may or may not be in the Shared Discourse Space). We now move on to discuss reference to entities in the speech situation. In this context, we are concerned with the demonstrative *that* (and *this*) rather than the complementizer. Here, too, we make a distinction between two types of reference: exophoric demonstratives referring directly to entities in the speech situation (discussed in this section) and anaphoric demonstratives referring to entities as represented in surrounding discourse (discussed in section 3.5).

Demonstratives are exophoric when they refer to entities "in the speech situation" (Diessel 1999: 93).[32] This is the prototypical use of demonstratives (e.g. *this*/*that book*) and can be accompanied by a pointing gesture. Traditionally, the distinction between the exophoric demonstratives *this* and *that* is taken to indicate the physical distance between the referent and the deictic origo (typically, the Speaker). However, a wealth of experimental studies have shown this view to be too simplistic (Peeters et al. 2021). Physical aspects of the relation between Speaker and referent are only one of a number of factors determining the choice of demonstrative. There are also psychological factors at play, which relate to "*the cognitive status of the referent* in the mind of the speaker and/or the addressee as assumed by the speaker" (Peeters et al. 2021: 412, emphasis original).[33] For example, different demonstratives may be chosen depending on whether the referent is in joint attention or whether it is considered cognitively accessible by the Addressee (Peeters et al. 2021: 413 and references therein).

Depending on context, different factors may weigh more or less heavily in the choice for a particular demonstrative. Peeters et al. (2021: 416–419) show how this works in Spanish, a language with a three-term distance contrast between *este* (proximal), *ese* (medial), and *aquel* (distal). In an experimental setting where a Speaker has to indicate one of a number of objects to an Addressee across the table, Coventry et al. (2008) found that *este* can only be used for objects in a relatively small zone around the Speaker, excluding most of the table and the Addressee on the other side. At first sight, this seems to be at odds with Jungbluth (2003), who showed that the range

32 In our view, anaphoric demonstratives also refer to entities in the speech situation, but only indirectly, via an antecedent in the surrounding discourse.

33 The choice of a demonstrative also depends on referent-intrinsic factors like animacy and grammatical gender, but these are not relevant to us here.

of *este* encompasses the entire conversational dyad, including both Speaker and Addressee. However, unlike Coventry et al. (2008), Jungbluth (2003) relies on natural data. Peeters et al. (2021) argue that psychological factors are not available in Coventry et al.'s (2008) experimental setting, prompting interlocutors to interpret the proximal/medial/distal distinction using physical factors like distance, and "calibrating" the different demonstratives to maximize information density. In natural language, however, psychological factors are more important, which explains the different results found by Jungbluth (2003).

In our analysis, psychological factors correspond to Addressee involvement, i.e. the recycling of the spatial relation between referent and Speaker to indicate whether the referent is "shared" with the Addressee. Entities are psychologically further from an interlocutor when they are not in attention or less accessible or identifiable. As above, we propose that English *that* refers to an element of the Shared Discourse Space, while *this* refers to an element in the Speaker's Personal Discourse Space. These psychological factors can be further interpreted pragmatically. Consider the following examples:

(3.15) a. *How's that throat?* (Lakoff 1974 in citation by Cheshire 1996: 376)

b. *How is that term paper coming along?*
(E. Riddle, p.c., in citation by Chen 1990: 150)

The demonstrative in (3.15a) could in principle be replaced by *your* or *the*. According to Cheshire (1996: 376), *your* would be unmarked, simply indicating awareness of the Addressee's illness, while *the* would make previous knowledge of the illness explicit. According to her analysis, *that* does not only signal this previous knowledge but also expresses Speaker involvement, which can be interpreted as empathy with the Addressee. Example (3.15b) can be analyzed analogously. Using our terminology, we could say that the Speaker uses *that* to signal that the throat is in the Addressee's and their own joint attention, and that this joint attention is what triggers the sympathetic reading. Lakoff (1974) and Cheshire (1996) do not discuss the interpretation of (3.15a) with *this*. This sentence seems quite unnatural, but we could imagine (*Let's see,*) *how's this throat?* in a context where a doctor begins to physically imagine a patient's throat. In this situation, the doctor is not interested in the patient's own judgment — and this corresponds to the lack of Addressee involvement marked by [–distal] *this*.

Kirsner (1979) discusses examples such as the following in Dutch:

(3.16) Dutch (Kirsner 1979: 357)

a. *Het is smoorheet, iedereen puft en bakt en in die/?deze hitte moet ik alles belopen.*
'It is boiling hot, we are all positively melting, and in *that*/?*this* heat I have to walk everywhere.'
(Anne Frank, 1959, *Het Achterhuis* [The diary of a young girl])

b. *"Ha die/*deze Frits!" zei de jongen, gaf hem een harde klap op de schouder, bleef voor hem staan en zei ...*
"'Aha, (*that*/**this*) Frits!," the boy said, slapped him on the shoulder, remained standing right in front of him and said ...'
(G. van het Reve, 1961, *De avonden* [The evenings])

In neither case can the use of the distal demonstrative be explained using physical distance: in (3.16a), the heat is immediately experienced by the Speaker, and in (3.16b), the Speaker must be close to the Addressee (given that he slaps him on the shoulder). Instead, Kirsner (1979) proposes that a proximal demonstrative would indicate that the Addressee must do relatively much work to identify the referent, compared to when the distal demonstrative is used. This is consistent with our notion of Addressee involvement: in our view, the referent of a distal demonstrative is already tracked by the Addressee, and would therefore require less work to identify.

A somewhat intuitive explanation for the contrasts in (3.16), which we do not support, relies on emotional distancing. Similar to Chen (1990) for other examples we might suggest that the use of a [+distal] demonstrative in (3.16a) creates distance between the Speaker and the referent, because the Speaker has a negative attitude towards the heat. Note, however, that in (3.16b) the [+distal] demonstrative is used in an intimate, amicable greeting. Chen (1990) suggests that *that* can express both emotional distancing and sympathy. But since these two are near polar opposites, this seems unlikely to us. We do not deny that *that* can be used in both positive and negative contexts, but we reject the analysis in which *that* can express both a positive and a negative attitude. Instead, *that* could express a more general notion, and the specific attitude could be derived from this general notion in conjunction with context. Cheshire (1996: 377) calls this notion "interpersonal involvement". We see it as an instance of Addressee involvement, since in both cases the Speaker is assuming shared context with the Addressee.

Let us then turn to the physical factors determining the choice of the demonstrative. In our model, these correspond to actual distance, i.e. the recycling of the spatial relation between the deictic expression and its referent. Note that exophoric demonstratives are often if not always accompanied by a pointing gesture, and can even be replaced by one (Jouitteau 2004: 431). We take this as an indication that the demonstrative has a position in the physical world, like the referent. Therefore, actual distance, that is, the relationship between the referent (the entity) and the deictic expression (the demonstrative), is determined by physical factors like Euclidean distance in the real world.[34]

3.5 Anaphoric reference and conversational interaction

Like exophoric demonstratives, anaphoric demonstratives refer to entities in the speech situation. However, they do so indirectly, by referring to a noun phrase in the surrounding discourse:

(3.17) German (Diessel 1999: 96)

[*Der*	*Anwalt*]$_i$	*sprach*	*mit*	[*einem*	*Klienten*]$_j$.	*Da*	*er*$_i$/*der*$_j$	*nicht*
the	lawyer	talked	with	a	client	since	he/this_one	not
viel	*Zeit*	*hatte,*	*vereinbarten*	*sie*	*ein*	*weiteres*	*Gespräch*	
much	time	had	agreed_on	they	a	further	conversation	
nächste	*Woche.*							
next	week							

'The lawyer talked to a client. Since he didn't have much time, they agreed to have another meeting next week.'

Unlike the personal pronoun *er*, the demonstrative pronoun *der* can only be coreferential with *ein Klient* 'a client': the demonstrative pronoun indicates a topic shift (Diessel 1999: 96). We also use the term anaphoric for demonstratives referring to (the interpretation of) larger bodies of text:[35]

34 Physical factors also include things like visibility, knownness, and elevation (Diessel 1999: 35–47; Peeters et al. 2021), but we focus on physical distance here.

35 This is part of what Diessel (1999: 100–105) calls the discourse deictic use of demonstratives. However, we only include references to propositions here (e.g. *That's false*), not references to illocutions (e.g. *That's a lie*). The latter are more like exophoric reference for us, since illocutions have properties like phonological form, which give them a place in the real world.

(3.18) [*Sales have been going up since 2019*]$_i$. [*This trend*]$_i$ *is the result of a growing interest...*

An intuitive hypothesis concerning the difference between *this* and *that* in these contexts would be that *this* refers to referents that are more proximal, in terms of either distance (length of text between antecedent and anaphor) or focus (*this* referring to newer or more important information; cf. Strauss 2002). Experimental work of Çokal et al. (2014) found no evidence for this, however, and other studies have found that proximal demonstratives are more likely than distal demonstratives to refer to antecedents further back in the text, contrary to what such an intuitive hypothesis would predict (Maes et al. 2022). Yet another problem for this intuitive hypothesis is that anaphoric *this* and *that* cannot be used contrastively (3.19b) while their exophoric counterparts can (3.19a):

(3.19) a. *I don't want this one, give me that one.* (distinguishing two objects on a table)

b. **I went Christmas shopping and bought a t-shirt*$_i$ *and a CD*$_j$*; that*$_i$ *is for Kim, and this*$_j$ *is for Paul.* (Stirling & Huddleston 2002: 1506)

All in all, there does not seem to be any positive evidence for exploitation of the actual distance, that is, properties of the relation between deictic expression and referent. We return to this issue in the conclusion.

However, the choice between a proximal and distal demonstrative does seem to be conditioned by Addressee involvement: the relations between the referent and the interlocutors. Evidence for this comes from corpus linguistics, in particular when it comes to the comparison of different corpora. According to Peeters et al. (2021: 421), the ratio of proximal vs. distal anaphoric demonstratives varies widely as a function of text or discourse genre. The strongest preference for proximal demonstratives is found in scientific, expository literature, whereas interactional spoken discourse shows a preference for distal demonstratives. Distal demonstratives are also preferred in written news stories, but to a lesser extent. Peeters et al. already recognize that the main difference between these types of corpora is the type of interaction between Speaker (writer) and Addressee ("news corpora ... in which information is clearly targeted towards the news item's consumer"; Peeters et al. 2021: 421). We can make this more concrete with the notion of Shared Discourse Space. In spoken dialogue, there is continuous feedback from the Addressee to the Speaker. As a result, the Speaker can be rel-

atively sure that the Addressee follows along and is attentively involved in the discourse. Thus, as with exophoric demonstratives, the use of the distal form here suggests reference to an element of the Shared Discourse Space between Speaker and Addressee. The same is true for news stories, which are written to be easily accessible by a wide audience. They are somewhat like monologues: there is no feedback from the Addressee, but the content is adjusted so that the Speaker can assume that the Addressee can follow. This is not true for scientific literature, where the high information density and wide variety of reader backgrounds seem to prevent the writer from assuming a large Shared Discourse Space with the Addressee. This means that scientific authors will use proximal demonstratives more frequently because they will typically assume that their readers do not share in their Personal Discourse Space.[36]

These hypotheses have been confirmed for written text in a corpus study by Maes et al. (2022), on the basis of written news stories, Wikipedia articles, and product reviews: "Text genres can be seen as carrying a default assumed psychological distance between writer and referents" (Maes et al. 2022: 26). An anonymous reviewer remarks that these correlations between genre and demonstrative variance can also be related to other factors, such as register (*that* being less formal). We agree that more work needs to be done in this area. However, at this point an explanation based on Addressee involvement strikes us as more economical. Addressee involvement can be related to the [±distal] feature that demonstratives obviously carry, and is independently needed to explain the data described in sections 3.2 to 3.4. Since the same notion can also explain the genre effect observed by Maes et al. (2022), there is, lacking evidence to the contrary, no need to overcomplicate things by adding a register feature to the analysis.[37] We conclude with

36 This suggestion generates falsifiable hypotheses that can be tested against other types of corpora. For instance, we would expect spoken monologues to show a slightly lower preference for distal demonstratives than interactional discourse, because there is less feedback from the Addressee. Also the fact that evaluative discourse shows a lesser preference for distal demonstratives than regular interactional discourse (Peeters et al. 2021: 421) can be explained this way, since evaluations are inherently personal and not in the Shared Discourse Space. On the other end of the spectrum we would expect to find more distal demonstratives in oral scientific discourse (e.g., conference presentations) than in scientific literature.

37 Note also that if we were to explain the genre effect with register, it is not clear yet why *that* would be associated with more informal registers, since there does not seem to be

Peeters et al. (2021: 422) that the choice between anaphoric *this* and *that* is conditioned primarily by the question whether the referent is "in close psychological proximity to the knowledgeable speaker or writer" or in "the shared space between speaker and addressee". The proximal/distal distinction in anaphoric demonstratives is therefore primarily recycled to mark Addressee involvement.

3.6 Conclusion

3.6.1 Generalizing over sentential and nominal reference

We have proposed a unified analysis of the recycling of the proximal/distal distinction between the demonstratives *this* and *that* in terms of both *actual distance* (the "distance" between deictic expression and referent) and *Addressee involvement* (the "distance" between Speaker and referent). This theory is also able to explain the correlation of *this* and *that* with direct and indirect speech reports, respectively, as well as the alternation between *that* (or a parallel finite complementizer) and a zero complementizer in a variety of contexts. These abstract distances are interpreted in different ways depending on the type of referent, as shown in table 3.1, reproduced here from the introduction.

The four types of environments discussed above have been categorized according to two binary properties here. First, our deictic elements refer to either information content or entities. We studied information content in sections 3.2 (direct and indirect speech) and 3.3 (presuppositions), and entities in sections 3.4 and 3.5 (exophoric and anaphoric demonstratives, respectively). Second, the well-known distinction between exophoric and anaphoric demonstratives for reference to entities generalizes to information content, where it distinguishes utterances from their meaning. Both exophoric demonstratives and speech reports refer directly to concrete things in the world (entities and utterances), whereas anaphoric demonstratives and the complementizers referring to the Shared Discourse Space refer only indirectly (to entities via linguistic antecedents, and to information content through a mental model of the discourse state).

anything [+distal] about informality (we might expect the contrary!). However, if we have an independent explanation for the genre effect based on Addressee involvement, the correlation with register is a simple consequence of the genre effect.

	Entities (DP)	**Information content (CP)**
Exophoric	Exophoric demonstratives (section 3.4):	Direct and indirect speech (section 3.2):
	Actual distance in the concrete physical world	Actual distance in a multidimensional conceptual world, interpreted as similarity
	Addressee involvement: interpreted as psychological factors (psychological distance, joint attention, empathy, ...)	Addressee involvement: interpreted as evidentiality; proximity is private witness evidentiality
Anaphoric	Anaphoric demonstratives (section 3.5):	Presupposition (section 3.3):
	Addressee involvement: *that* used over *this* to interact and empathize with the Addressee	Addressee involvement: *that* used over zero to signal content in the Shared Discourse Space

Table 3.1 Deriving different kinds of reference from two binary properties.

There are two gaps in table 3.1. First, it becomes clear that actual distance (the "distance" between referent and deictic expression) is not used in anaphoric reference. We can understand why this is the case in the following way. In both exophoric and anaphoric reference there is a direct link between the referent and the Speaker, namely in the cognitive model of the Speaker. This allows the proximal/distal distinction to be recycled to mark Addressee involvement. But a direct link between the referent and the deictic expression, which is needed to describe actual distance, only exists in exophoric reference: in anaphoric reference, the link is indirect, through an intermediate linguistic entity. The fact that this link is indirect seems to make it difficult to interpret the distance expressed by the proximal/distal element in terms of the relation between referent and deictic expression in these cases, and therefore there is no actual distance there.

The second gap is that proximal elements appear to be incompatible with anaphoric reference to information content (presuppositions): complementizer *that* alternates with a zero complementizer rather than with a complementizer based on proximal *this*. This gap has already been explained in section 3.3.4: using a [–distal] element in this type of reference would suggest that the Speaker refers to informational content that is new

to the Addressee (because it is not in the Shared Discourse Space) without introducing it (because anaphoric reference is used). Such use of language would be incompatible with cooperative conversation. In other words, the analysis presented here explains why finite complementizers are so rarely derived from proximal demonstratives.

What unifies the interpretation of Addressee involvement in all four contexts is the fact that the referent is presented as accessible to, or tracked by, the Addressee. Depending on the context, this may have some further implications. This is particularly visible with demonstrative *that*, as shown in section 3.4. In these contexts, *that* is in opposition not only with *this* (which would explicitly mark the referent as in the Personal Discourse Space of the Speaker) but also with the definite article *the*. The latter has no [±distal] feature, but can still be used in contexts where the referent is mentally accessible to the Addressee. Consider the following contrast:

(3.20) a. *Could you pass me the hammer?*
b. *Could you pass me that hammer?*

Example (3.20a) can be used in a context where the Addressee is either already tracking the hammer in their Personal Discourse Space, or can easily identify it — that is, *the* already implies Addressee involvement. As a result, the meaning of *that* becomes more marked: excluding pointing contexts where actual distance is promoted, (3.20b) is most natural in situations where the Addressee is already tracking the hammer, not in situations where the hammer is only identifiable. We thus see that the interpretation of the Addressee involvement marked by [+distal] *that* becomes more marked when it enters into an opposition with *the*. Such an opposition is not available for the complementizer *that*. As a result, the interpretation of Addressee involvement is simpler in this environment, and is confined to referring to an element of the Shared Discourse Space.

3.6.2 Related work

In this chapter we have sought to bring together a number of well-known and much studied phenomena in a single theory. We do not have space here to review the full history of scholarship of all these phenomena individually. However, work on some of these issues has, without relating them to the other phenomena, reached similar conclusions to ours, and therefore deserves discussion here.

In particular, there is a long history of work on so-called optional *that* in English, some of which has been referred to in section 3.3.2. First of all, Yaguchi (2001) presents an analysis that is quite close to ours: her paper "elucidates the function of the non-deictic *that* by considering how the residual meaning of the demonstrative *that* is still in effect [...] and what underpins the presence or absence of the non-deictic *that* from a cognitive perspective" (Yaguchi 2001: 1126). While Yaguchi reaches similar descriptive generalizations based on similar data, we believe the analysis needs refinement. In particular, Yaguchi (2001: 1127) describes the complementizer *that* as "non-deictic", while she claims at the same time that it also preserves the function of the demonstrative to "deictically point". It is unclear how the two can be reconciled. Furthermore, Yaguchi (2001: 1127) takes a leap by assuming that "non-deictic" *that* has to do with truth: "the use of demonstratives implicitly encodes the speaker's presupposition that the hearer can identify the entity to which the speaker refers [...] By the same token, non-deictic *that* [...] signals that the speaker presupposes the contents of the complement clause to be referential, in other words, to contain true or valid information, whose validity can be proven by evidence." We agree that *that* is referential in both uses and that this can entail presupposition, but stress that it is perfectly possible to refer to things that are not true or valid. This shows, for instance, in (3.7) above, where *that* is used to acknowledge an implicit question of the Addressee:

(3.7) a. *I thought you might need some help.* (Bolinger 1972: 58)

b. *I thought that you might need some help.* (Bolinger 1972: 58)

For this reason we have analyzed *that* using references to a *Shared Discourse Space* which, unlike the Common Ground, does not only contain presupposed propositions but also other information content, such as questions or rejected propositions, and entities. Yaguchi (2001: 1137–1139) discusses verbs like *think*, *believe*, and *guess*, which do not presuppose their complement, but does not use referentiality to explain the use of *that* with these verbs. Instead, the distance expressed by *that* would mark the greater amount of evidence and analytical thinking used to come to the conclusion stated in the complement. Yaguchi does not specify, however, how speakers choose between these different factors (referentiality and amount of evidence) when interpreting an instance of *that*. Furthermore, Yaguchi's approach is problematic for verbs like *doubt*, which suggest that the Speaker

favors presupposing the negation of the complement. Even if the use of *that* in *I doubt that P* has to do with the amount of evidence, it has to do with the amount of evidence for ¬*P* rather than for *P*. By contrast, in our account we analyze all these different cases as involving referentiality. For example, the use of *that* in *I doubt that P* reflects that the question whether *P* is the case is tracked by both Speaker and Addressee.

Dor's (2005) position on optional *that* is also quite similar to ours: he suggests that "the predicates which can embed the bare clause, without the complementizer, are those which entail that a cognitive agent (in the majority of cases, their subject) has made an epistemic claim concerning the truth of the proposition denoted by the embedded clause" (Dor 2005: 347). This improves on Yaguchi (2001) since it accounts for verbs like *doubt*, though we would still widen the scope a bit to involve reference to questions (for which no truth claim has been made) as well. Furthermore, Dor (2005) is primarily descriptive and does not seek to explain why the use of *that* is related to truth claims. In our view (as in that of Bolinger 1972 and Yaguchi 2001) this can be explained as a type of reference, and thus connected to the demonstrative *that*.

This brings us to related work on the similarities between demonstratives and finite complementizers highlighted in table 3.1 above. Roberts & Roussou (2003: 111–116) dismantle a number of arguments for the supposed synchronic homophony of demonstrative and complementizer *that*, which is the basis for much of what we are doing here. Kayne (2014) argues that the complementizer *that* is still a demonstrative, but one that does not require "pointing". He also addresses the question why *this* is not a complementizer, providing an explanation based on a first person feature as opposed to our [±distal]. This is compatible with our analysis if first person is seen as an interpretation of [–distal]. Most recently, Ritter & Wiltschko (2019) and Colasanti & Wiltschko (2019) have argued for a nominal Speech Act structure dominating the DP layer. As is well known, Speech Act structure on the CP level is used to mark the relationships between the propositional content and the Speaker and Addressee, thus formalizing the differences between declaratives, exclamatives, interrogatives, and other sentence types. On the DP level, the Speech Act structure would be used to express the relationships of the interlocutors and the described entity — in particular whether it is discourse-old or discourse-new. This formalization is readily applicable to the observations we have discussed in this chapter.

3.6.3 Final remarks

By way of conclusion we want to discuss three final points. First, we wish to point out that paying attention to the fact that the two abstract distances are recycled in different ways depending on the type of reference allows us to resolve some apparent paradoxes. For instance, recall that Cheshire (1996) argued that the exophoric demonstrative *that* can express empathy with the Addressee:

(3.15a) *How's that throat?* (Lakoff 1974 in citation by Cheshire 1996: 376)

On the other hand, Storms (1966) suggested that in the context of a witness interrogation, sentences without *that* are used "to put the witness at her ease and at the same time to set an unsuspected trap" (Storms 1966: 263). Thus, the demonstrative *that* in (3.15a) would engage with the Addressee, whereas it is the absence of the complementizer that does this for Storms (1966). By fleshing out what Addressee involvement really means in these different types of environments, the apparent paradox can be resolved: Cheshire (1996) is talking about reference to entities, where the distal demonstrative establishes joint attention and hence empathy; Storms (1966) is talking about information content where Addressee involvement concerns the Common Ground, and hence the establishment of facts. In this way, Addressee involvement is a useful generalization from which other notions, such as empathy (Cheshire 1996) or "relating to knowledge" (Wierzbicka 1988), can be derived.

Second, a unified analysis of demonstratives and complementizers allows us to explain why *that* introduces finite complements rather than non-finite ones. Tsoulas (1996: 298) points out that the finite/non-finite distinction in clausal complementation can be better described in terms of "definite" and "indefinite" propositions. A proposition is definite when it uses a "definite" tense, that is, a tense that specifies a precise temporal point. In this sense, finite complements are "definite" and infinitival complements are "indefinite"; the latter can by their nature not be situated precisely in space. The selection of a tensed complement by the complementizer *that* can be derived from its demonstrative nature: it references the precise temporal point. In other words, the fact that the complementizer *that* takes finite complements is fully analogous to the fact that demonstratives are necessarily definite (in the common sense): both require their referent to be situated in space and time.

Finally, one might wonder where the relativizer *that* fits in table 3.1. Its position is clearly in the lower left quadrant for anaphoric reference to entities. However, note that there is no [–distal] relativizer (*the book *this*/*that is on the table here*), which matches with the complementizer *that* in the lower right quadrant (anaphoric reference to information content). We can explain the lack of a proximal relativizer in the same way as we explained the lack of proximal reference to information content: since the referent/antecedent is mentioned in the immediately surrounding context, it is necessarily in the Shared Discourse Space and can therefore not be referred to by a proximal element. Therefore, although the relativizer *that* stands in the lower left quadrant, Addressee involvement is interpreted not as interaction/empathy with the Addressee (as with other anaphoric reference to entities) but using Shared Discourse Space (as with reference to information content). We thus find the distinction between overt and zero relativizers to be similar to that between overt and zero complementizers. For example, (3.21a) is uttered out of the blue by a detective sergeant to a responding officer, and the Speaker does not expect there to have been anything unusual. The relative clause thus does not have any grounding in space-time or previous discourse, and *that* can be omitted. On the other hand, suppose a customer is looking through the racks in a clothing store. The salesclerk may then ask (3.21b), where a zero complementizer would be odd: the fact that the customer is looking for something is presupposed. There is a well-defined set of items from which the answer can be drawn (all the clothes in the racks), in contrast to the open-ended nature of (3.21a).

(3.21) a. *There was nothing unusual Ø caught your eye when you came in?*
(*Inspector Morse*, season 7, episode 1)

b. *Was there anything that*/*?Ø caught your eye while browsing through the racks?*

In this chapter we have analyzed a number of high-frequency uses of the proximal/distal distinction, but our discussion has not been comprehensive. It is expected that actual distance and Addressee involvement can be interpreted differently in other contexts. What we do commit to is the position that the proximal/distal distinction is interpreted in terms of the distance between Speaker and referent (Addressee involvement) and/or deictic expression and referent (actual distance). In this way, this chapter provides an instrumentarium for further analysis of other kinds of reference.

Biblical Hebrew *kī* 'that' as a marker of Common Ground

4

Abstract The Biblical Hebrew clausal connective כִּי *kī* has many functions: introducing object and subject clauses as well as causal, temporal, conditional, adversative, concessive, and resultative adverbials. How are these functions related, and can they can be reduced to a single semantic core? In this chapter I describe *kī* as a marker of Common Ground. This imposes significant constraints on the contexts in which *kī* can be expected, and distinguishes *kī* from other words with similar functions. The chapter also explains why *kī* should mark Common Ground: this follows from the recycling of an original [+distal] feature of the Proto-Semitic morpheme **ka*. I further argue that most of the different adverbial readings of *kī* are not the result of lexicalization, but rather pragmatically inferred, based on the general notion of Common Ground together with context-specific information. This provides a more economical description of this problematic lexeme.

4.1 Introduction

In the previous chapter it was argued that the presence or absence of "neutral" complementizers like English *that* can depend on the information status of the content of the complement clause: *that* is more often used when the information in the complement clause is in the Common Ground. This is most easily seen in exclamatives:

(4.1) ***That** bio industry is still allowed!* (= [3.3a])

Unlike a declarative sentence without *that*, (4.1) cannot be used to convince the Addressee of the fact that bio industry is still allowed. In the analysis of chapter 3, the presuppositional status of the propositional content is marked by the presence of complementizer *that*.

This chapter is first published here. I am grateful to Ellen van Wolde and Johan Rooryck for useful feedback on this chapter. Any remaining mistakes are mine alone.

Other clausal connectives are also sensitive to information status. As just one example, English *for* is typically used to introduce parenthetical causal clauses, which contain backgrounded information that may already be familiar to the Addressee. Thus, speakers for whom *for* is still productive prefer *for* over *because* in (4.2a), while *because* is preferred in (4.2b), where new information is introduced.[1] In sentences like (4.2a), the Speaker assumes that the information in the *for*-clause is known, or can readily be assumed by, the Addressee.[2]

(4.2) a. *An automatic timer would soon turn* [*the light*] *off,* ***for*** *we* [*Ladover Jews*] *do not tamper with electricity on Shabbos.*
(Chaim Potok, 1990, *The gift of Asher Lev*)

b. *My mother's sister ..., who had been unable to attend the funeral* ***because*** *her husband had undergone bypass surgery ..., flew in from Boston.*
(Chaim Potok, 1990, *The gift of Asher Lev*)

In this chapter I extend the analysis of chapter 3 to a highly polysemous clausal connective in an unrelated language to demonstrate the wide applicability of the theory. I work out the case of the Biblical Hebrew clausal connective כִּי *kī* and illustrate the different discursive effects a reference to the Common Ground can have. I will argue below that like English *that*, *kī* carries a [+distal] feature, which makes a Common Ground analysis a priori likely. A look at any dictionary suggests a plethora of different uses for *kī*: introducing object and subject clauses ('that'), causal 'because, for', temporal 'when', conditional 'if', adversative 'but', concessive 'though', resultative 'so that', and more. Previous scholarship has failed to reduce these different uses to a single semantic core. I argue that marking Common Ground could constitute this semantic core, and that the different uses can be derived from syntactic and pragmatic clues based on this general semantics.

The chapter thus makes two contributions: it shows that clausal connectives in unrelated languages are sensitive to reference to Common Ground, and it discusses in depth the different discursive effects reference to Common Ground can have. The remainder of the introduction summarizes the framework of chapter 3 which forms the basis for this analysis (section 4.1.1),

1 Similarly, *since* introduces specifically not-at-issue causal clauses compared to *because* (Charnavel 2017). There is some correlation between discourse-old information status and not-at-issueness, since discourse-new information content is typically at-issue.

2 *For*-clauses can be analyzed as right dislocations (De Vos in preparation), which are associated with discourse-old or inferential information status (e.g. Grosz & Ziv 1998).

and also provides a brief history of scholarship on *kī* (section 4.1.2). In section 4.2 I describe my method and provide an overview of the data. The following sections present an in depth analysis of the different ways in which *kī* can be used, in which I show how each use derives from the core function of marking Common Ground (sections 4.3 to 4.7). Section 4.8 concludes.

4.1.1 Theoretical framework

This subsection briefly summarizes the framework developed in chapter 3. Readers who have read chapter 3 can safely skip ahead to section 4.1.2.

One of the main claims of chapter 3 was that the [+distal] feature of the English demonstrative *that* is still present, but differently interpreted, when *that* is used as a complementizer. English *that* marks not only spatial distance (**this*/*that book over there*) but also involvement with the Addressee, who is "far" from the Speaker. This notion is interpreted in different ways depending on the context; in the sentential domain, complementizer *that* refers to Common Ground between Speaker and Addressee.[3] In particular, the Common Ground is conceptualized as accessible and "close" to the Addressee and, as a result, "far" from the Speaker.

Figure 4.1 (reproduced from figure 3.1) clarifies the model. The circle around S stands for the information content tracked by the Speaker; the circle around A for the information content tracked by the Addressee. The intersection of these sets (shaded dark gray) can, under some assumptions, be seen as the Common Ground.[4] The light gray shaded region represents the Speaker-private information content; information that is not tracked by the Addressee. Even though both gray regions are equally accessible to the Speaker, languages use [−distal] and [+distal] forms to refer to the light and the dark region, respectively. Thus, while the distance is measured from the Speaker's origo, it reflects the accessibility to the Addressee, i.e., Common Ground status.

3 Chapter 3 introduced the term "Shared Discourse Space" for the region tracked by both the Speaker and the Addressee. Shared Discourse Space is more general than Common Ground and does not only contain information content. Since this chapter deals exclusively with the sentential domain, the more common term "Common Ground" suffices here. The argument is fully compatible with an analysis based on Shared Discourse Space, should this be needed for data not covered in this study.

4 Most importantly, the circle around the Addressee represents what *the Speaker assumes to be* tracked by the Addressee (see section 3.2).

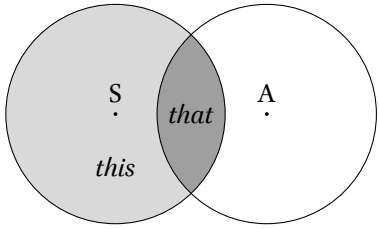


Figure 4.1 The information content tracked by the Speaker and Addressee. The intersection, the Common Ground, is seen as "far" from the Speaker.

Most notably, Addressee involvement plays a role in the interpretation of complementizer *that*. Since *that* is [+distal], it refers to the discourse-old information content in the Shared Discourse Space. This explains why *that* is used in exclamatives, whose propositional content is presupposed (see [4.1] above), as well as in other contexts discussed in chapter 3.

Speakers can interact with the spatial model in figure 4.1 in different ways. In the most basic cases, [+distal] forms are used to refer to discourse-old information content (which is in the Common Ground), and [−distal] forms are used to refer to discourse-new information content. However, there are two cases in which discourse-new information content can be presented as part of the Common Ground: when the content can be *accommodated* by the Addressee and when it is *imposed* by the Speaker. Declarative sentences with discourse-new information content are typically modeled as a request or proposal to update the Common Ground (cf. Farkas & Bruce 2010: 92). However, by explicitly placing the content "near" the Addressee with a [+distal] element, the Speaker can present it as if it is already part of the Common Ground. The Speaker does so to signal that they expect that the request for a Common Ground update will be granted, or, alternatively, to emphasize that they do not permit the Addressee to reject the proposal to update the Common Ground. In the first case, the Speaker assumes that the Addressee can *accommodate* the information content; in the second case, the Speaker *imposes* the content on the Common Ground.[5]

These three types of reference (to discourse-old, accommodated, and imposed information content) can all be seen as referencing information

5 For more details and references, see Kocher (2022: 176–177). Some examples may be helpful here. The reader may browse ahead and compare cases of accommodation (e.g. [4.4; 4.17; 4.20; 4.23–4.24]) with those of imposition (e.g. [4.11; 4.19; 4.25]).

conceptually "near" the Addressee. Discourse-old information content is "near" the Addressee because it is known and accessible to her. When new information content is placed "near" the Addressee by the Speaker, she can thereby suggest that the Addressee should easily be able to accommodate it. When this is not the case, this forces the Addressee to react; this is a case of information content imposed on the Common Ground. It is not surprising, then, that an originally [+distal] deictic element like English *that* can be used to interact with the Common Ground in these different ways: [+distal] *that* refers to a space "far" from the Speaker, but "near" the Addressee.

4.1.2 Biblical Hebrew *kī*

As mentioned above, Biblical Hebrew *kī* has many different uses: introducing object and subject clauses ('that'), causal 'because, for', temporal 'when', conditional 'if', adversative 'but', concessive 'though', resultative 'so that', and more.[6] It is generally accepted that *kī* derives from an originally deictic morpheme **ka*, so we are not dealing with multiple, accidentally homonymous particles. The morpheme **ka* is clearly ancient, given its appearance in at least Phoenician, Aramaic, and Arabic with similar functions (Lipiński 2001: §49.9). Many authors have pointed to this morpheme as evidence for a general "deictic" meaning of *kī* (e.g. Muilenburg 1961; Schoors 1981). What has not been given attention, however, is that **ka* is specifically a [+distal] deictic morpheme, referring to things at some distance from the Speaker.[7] This

6 Note that the distinction between main and subordinate clauses is not as strict in Hebrew as it is in, for example, English. For simplicity I will sometimes refer to *kī*-clauses as "subordinate" to a corresponding "main" clause, but it should be kept in mind that the relation between the two clauses is often more paratactic than hypotactic.

7 Lipiński (2001: §36.35, 36.37, 36.41) lists a handful of demonstratives in Semitic and beyond where **ka* appears to be proximal, but these are only a handful of isolated instances. Distal demonstratives take **ka* more often and more consistently. This is especially clear in West Semitic, where **ka* also appears in demonstratives. According to Hasselbach (2007: 3), **ka* in demonstratives "regularly marks far deixis in those languages in which it occurs". In some languages where demonstratives going back to **ka* are in paradigmatic contrast with the third person personal pronoun used as a [+distal] demonstrative, it appears that the forms based on **ka* are specifically used to refer to something near or known to the Addressee (e.g., 'give me *that* (from **ka*)' referring to an object near the Addressee in Jewish Babylonian Aramaic; Bar-Asher Siegal 2013: 82). This would align with the notion of Addressee involvement, but a discussion of these demonstrative forms is out of scope here.

[+distal] feature forms the basis for an interpretation based on Addressee involvement within the framework of chapter 3, summarized above.

There is no consensus as to how the different meanings of *kī* are related to each other and to original **ka*. In very broad strokes, the literature can be divided into those scholars who claim that all (or at least most) uses of *kī* can be reduced to a single semantic core (e.g. Muilenburg 1961; Schoors 1981), and those scholars who claim that diachronic processes like grammaticalization have lead to a highly polysemous lexeme (e.g. Locatell 2017, 2020).[8] The synchronic approach has been abandoned by most recent authors with the exception of Follingstad (2001), because the ways in which functions of *kī* can be said to be "deictic" are not well-defined, so that the theory is not constrained enough. Furthermore, it is unclear how some functions, like the causal one, can be reduced to the very general notion of deixis. On the other hand, diachronic developments can often be made conceivable but not proven. A diachronic account also does not answer the question how speakers could have understood which function of *kī* is used in a particular instance, given the high degree of polysemy.

I argue that a primarily synchronic account is possible using the notion of Addressee involvement introduced above.[9] I will show, for example, that *kī* is not used to introduce just any object clause, but specifically those object clauses whose information content is in the Common Ground. The same goes for other uses: *kī* cannot introduce just any temporal, conditional, or adversative (etc.) clause, but only those whose information content is in the Common Ground. This provides a more economical description of *kī* than a diachronic approach: roughly, *kī* is used when the information content of the complement clause is in the Common Ground. The particular interpretation (as causal, temporal, etc.), largely depends on context.[10]

8 See further Redslob (1835), Vriezen (1958), Muilenburg (1961), Schoors (1981), Bandstra (1982), Claassen (1983), Thorion (1984), Aejmelaeus (1986), Gross (1991), Benigni (1999), Follingstad (2001), Park (2016), and Locatell (2017, 2020). Some passages in works with a broader scope are relevant as well, in particular Watts (1964: 118–149); Muraoka (1985: 158–164); Van der Merwe (1993: 38–41); and Conklin (2011: 46–59). I will not summarize related work here, as this has recently been done by Locatell (2017), and only refer to these earlier sources where relevant to my own argument.

9 In two contexts I do need to assume a semantic shift, but in both cases the shift is typologically plausible.

10 One occasionally finds generic arguments against such a reductionist approach (Aejmelaeus 1986: 195; Locatell 2017: 114). However, the persistence of lexical meaning is,

4.2 Method and overview of the data

My analysis is based on an exhaustive analysis of the 808 uses of *kī* with a clausal complement in the narrative portions of the biblical books Genesis, Judges, Samuel, and Ruth. These books are considered to be similar in terms of time and place of origin, and thus form a relatively homogeneous corpus. I focus on narrative texts because these contain most direct speech, where we can expect the interaction with Common Ground to be the largest.[11] However, I also included uses of *kī* outside direct speech, where the narrator is the "Speaker" and the reader is the Addressee.

Each instance of *kī* was classified as belonging to one use type.[12] My classification of each instance can be found in the data set accompanying this chapter (Staps 2023b). The types are based on categories commonly found in reference works and literature on *kī*: (a) introducing object and subject clauses ('that'; tagged as "complementizer"), (b) causal 'because, for', (c) adversative 'but', (d) causal-adversative 'not X, *because*/*but rather* Y', (e) conditional 'if', (f) temporal 'when', (g) resultative 'so that', and (h) concessive 'though'.[13] Instances where the *kī*-clause does not seem to relate to a corresponding "main" clause were classified as (i) standalone; this group will be further subcategorized in section 4.7. Five cases were ambiguous; I will mostly ignore these for ease of exposition.[14] The distribution over the vari-

in fact, expected in grammaticalization processes (Hopper 1991: 28–30). It is therefore not surprising if aspects of the [+distal] deictic meaning of **ka* are preserved in *kī*, and my claim is that this is the most economical description of the data.

11 It is conceivable that the use of *kī* in poetry follows a different, but comparable, distribution (see e.g. Meyer 2001). In poetry it is often less clear what the Common Ground contains, so this may be a weaker factor in choosing between *kī* and alternatives in poetic texts. It is also possible that there are differences in distribution between narrative (the Speaker is the author) and direct speech reports (the Speaker is a character in the text). I will have to leave both questions for further study, however.

12 I excluded instances of the fossilized construction כִּי אִם *kī 'im*, and two instances of bare *kī*, in the meaning 'except', which I assume to have grammaticalized independently.

13 Naturally, there are cases that can be classified as one of two categories, in particular in the temporal/causal, temporal/conditional, and causal/resultative categories (cf. Locatell 2020). My argument does not depend on a sharp distinction between these categories, so I have in these cases selected what seemed to be the most relevant category without spending too much thought on it.

14 Gen. 8:21 (causal/concessive); 21:7 (adversative/standalone); 38:16c (conditional/resultative); 1 Sam. 15:24a (resultative/standalone); 2 Sam. 18:3b (causal *or* adversative, but not causal-adversative). See the data set for more details.

ous types is shown in figure 4.2. More than half of the occurrences are causal, and approximately one in four instances of *kī* introduces an object or subject clause.

Each instance was also tagged for the way it interacts with Common Ground. As explained in section 4.1.1, the information content of the *kī*-clause can be (a) discourse-old and thus part of the Common Ground, (b) easily accommodated by the Addressee as new Common Ground, (c) imposed on the Common Ground by the Speaker for discursive effect. There are some cases where the information content in the *kī*-clause does not fit either of these cases; these were classified as (d) rest. I will discuss these separately below, to show why they do not constitute counter-examples for my claim that *kī* marks reference to Common Ground. Nevertheless I also include many examples where information *is* in the Common Ground, in order to illustrate the various discursive effects this can have.

Figure 4.3 shows the distribution of types of reference to the Common Ground for each use type of *kī*; the numbers from which this graph has been compiled are given in table 4.1. It can immediately be seen that in the vast majority of cases, the information content of the *kī*-clause is discourse-old, easily accommodated, or imposed: in only 7% of the total number of cases there is no reference to Common Ground. The cases where *kī* apparently does not interact with Common Ground are mostly isolated in a few use types (causal, adversative, and causal-adversative).

Type	**Discourse-old**		**Accommodated**		**Imposed**		**Rest**	
Complementizer	150	75%	29	15%	18	9%	2	1%
Causal	245	57%	124	29%	23	5%	41	9%
Adversative	7	44%	3	19%	1	6%	5	31%
Causal-adversative	15	60%	5	20%	1	4%	16	25%
Conditional	4	57%	3	43%	0		0	
Temporal	15	50%	15	50%	0		0	
Resultative	35	85%	3	7%	3	7%	0	
Concessive	4	100%	0		0		0	
Standalone	15	47%	4	13%	12	38%	1	3%
Total	**490**	**62%**	**186**	**24%**	**58**	**7%**	**53**	**7%**

Table 4.1 Distribution of instances of *kī* in the data set, and the use of Common Ground per type.

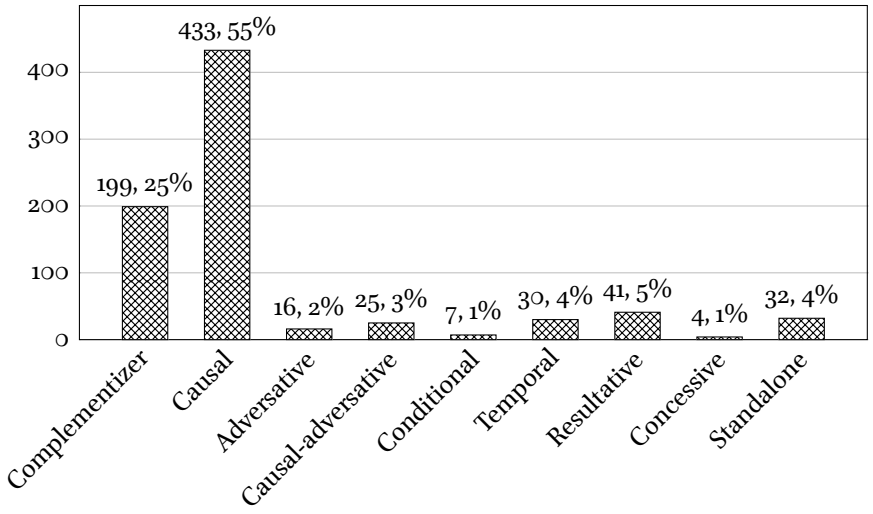


Figure 4.2 Distribution of *kī* over use types.

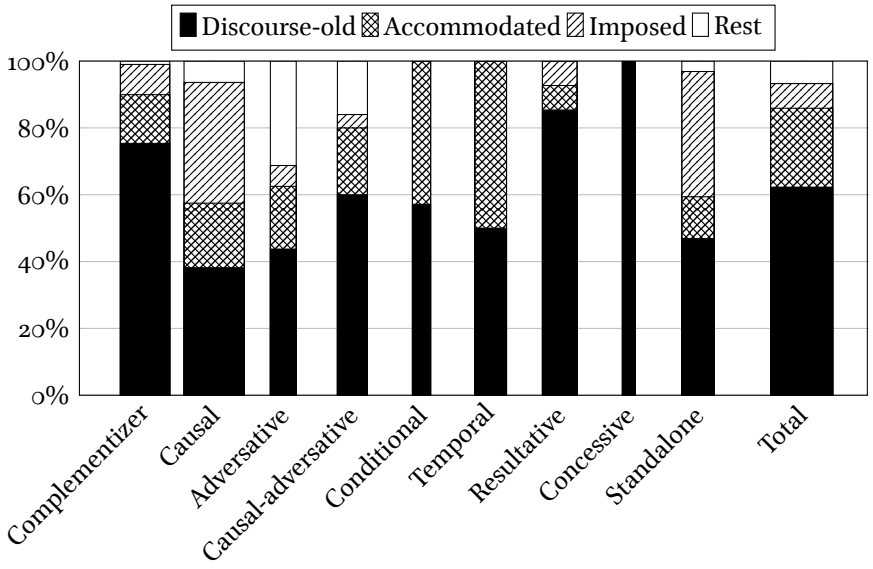


Figure 4.3 Distribution of types of reference to the Common Ground for each use type of *kī*. Column width indicates type frequency (not to scale).

4.3 Object and subject clauses introduced with *kī*

This section concerns cases of *kī* classified as "complementizer", in which *kī* is used to introduce an object or subject clause. It is cross-linguistically common for complementizers such as *kī* to develop from deictic elements.[15] I will not concern myself here with the question why an originally [+distal] deictic element becomes a complementizer; this has been addressed in great detail elsewhere.[16] Rather, I will compare complement clauses introduced by *kī* with other complementation strategies to show that *kī*-clauses are used specifically when reference to Common Ground is being made.

Object clauses introduced by *kī* can be divided into four categories depending on the type and semantic field of the matrix predicate. By far the most common are (a) verbs of perception (רָאָה *rāʾā* 'see'; שָׁמַע *šāmaʿ* 'hear') and (b) cognitive verbs (יָדַע *yāḏaʿ* 'know'); less common are (c) speech verbs (הִגִּיד *higgīḏ* 'inform') (see Miller 2003: 98 on the relative rarity of this category). The last category contains (d) miscellaneous constructions: cases where the object clause is governed by a noun (עֵד *ʿeḏ* 'witness') or preposition (עַד *ʿaḏ* 'until'), as well as cases where *kī* introduces a subject clause.

4.3.1 Verbs of perception

The verb רָאָה *rāʾā* 'see' occurs 65 times with *kī* in my corpus, שָׁמַע *šāmaʿ* 'hear' 20 times, and הִתְבַּשֵּׂר *hiṯbaśśer* 'receive good news' once. These verbs are most frequently used to describe an event in which the subject becomes

15 For example, English *that* is both a distal demonstrative (*that book*) and a complementizer (*I know that …*); it shares these functions with Semitic **ka*. In Latin, the complementizer *quod* (*Scio quod …* 'I know that …') is made up of the interrogative element *qu-* and the medial demonstrative *-id*. In Russian, the complementizer *što* (*ya znayu što …* 'I know that …') is related to the demonstrative *eto* (*eto dom* 'this house').

16 It is usually assumed that the English complementizer *that* developed from a cataphoric pronoun: *Galileo said that*$_i$*:* [*the earth is round*]$_i$ > *Galileo said* [CP *that the earth is round*] (Roberts & Roussou 2003: 113, and references therein). This grammaticalization path is somewhat problematic for Hebrew *kī*, which never was a demonstrative. However, this grammaticalization path has been challenged in recent work (Kayne 2014: 189; Axel-Tober 2017), for example because it does not explain why complementizers are typically based on [+distal] elements (**Galileo said* ***this*** *the earth is round*). For this reason, I assume that complementizers are not simply reanalyzed demonstratives, but lexicalizations of the same deictic [+distal] feature in a different syntactic environment (C rather than D); see section 3.3.4.

aware of information that was already known to the Addressee (usually, the reader of the text). It is better understood as 'realize' than as 'see', as in (4.3), where the information is clearly discourse-old, and therefore in the Common Ground for the Addressee (the reader).

(4.3) Gen. 16:4: וַיָּבֹא אֶל־הָגָר וַתַּהַר וַתֵּרֶא כִּי הָרָתָה

way-y-ăḇō'-Ø *'el* *hăḡār* *wat-t-ahar-Ø* *wat-t-ērɛ'-Ø*
WAYQ-3M-come-SG to Hagar WAYQ-F-conceive-3SG WAYQ-F-see-3SG

kī *hărăṯ-ṯā*
COMP conceive\PFV-3F.SG

'And he (Abraham) came into Hagar and she conceived, and she saw *that* she had conceived.'

The information can also be accommodated. In (4.4), the fact that the man cannot overpower Jacob can be accommodated by the Addressee given the information that they wrestle for a long time (until daybreak).

(4.4) Gen. 32:25–26: וַיֵּאָבֵק אִישׁ עִמּוֹ עַד עֲלוֹת הַשָּׁחַר: וַיַּרְא כִּי לֹא יָכֹל לוֹ

way-y-ē'ăḇēq-Ø *'îš* *'imm-ō* *'aḏ* *'ălōṯ* *haš=šāḥar*
WAYQ-3M-wrestle-SG man with-him until go_up\INF.of the=daybreak

way-y-ar'-Ø ***kī*** *lō'* *yāḵōl-Ø* *l-ō*
WAYQ-3M-see-SG **COMP** not be_able\PFV-3M.SG to-him

'And a man wrestled with him (Jacob) until daybreak, and he realized *that* he could not overpower him.'

A comparison with other complementation strategies is most helpful to show that *kī*-clauses are associated with Common Ground; I will use רָאָה *rāʾâ* 'see' as a running example. This verb occurs with three types of clausal complements in my corpus: besides *kī*, an interrogative pronoun can be used, or the complement can be introduced with וְהִנֵּה *wə-hinnē* 'and behold'. Unsurprisingly, complement clauses headed by an interrogative pronoun express a lack of information and are therefore incompatible with *kī*, which marks Common Ground:

(4.5) Gen. 37:20: וְעַתָּה | לְכוּ וְנַהַרְגֵהוּ ... וְנִרְאֶה מַה־יִּהְיוּ חֲלֹמֹתָיו:

wə='attā *lə-ḵ-ū* *wə=n-aharḡ-ēhū* ... *wə=n-ir'ɛ*
and=now go\IMP-M.PL and=1PL-kill\IPFV-him ... and=1PL-see\IPFV

mah *y-ihy-ū* *ḥălōm-ōṯ-āyw*
what 3M-be\IPFV-PL dream(M)-PL-his

'Now, let's kill him ... and we'll see *what* will become of his dreams!'

More interesting is the division of labor between *kī* and *hinnē* 'behold'. With the latter, the information given in the complement clause is new and not anticipated by the Addressee:[17]

(4.6) Gen. 22:13: וַיִּשָּׂא אַבְרָהָם אֶת־עֵינָיו וַיַּרְא וְהִנֵּה־אַיִל אַחַר נֶאֱחַז בַּסְּבַךְ בְּקַרְנָיו

way-y-iśśå'-Ø *'aḇrå̄hå̄m* *'ɛṯ* *'ēn-å̄yw* *way-y-ar'-Ø* ***wə=hinnē***
WAYQ-3M-lift-SG Abraham OBJ eye-DU.his WAYQ-3M-see-SG **and=behold**

'ayil *'aḥar* *nɛ'ĕḥaz-Ø* *b=as=səḇaḵ* *bə=qarn-å̄yw*
ram behind hold\MID.PFV-3M.SG in=the=bush in=horn-PL.its

'As Abraham looked up, he saw — *and behold!* — a ram behind[18] [him]; it had been caught with its horns in a bush.'

However, in some cases with *hinnē* 'behold', the information in the complement clause is not new to the Addressee (the reader), but new to the subject of *rå̄'å̄* 'see':[19]

(4.7) Gen. 8:13: חָרְבוּ הַמַּיִם מֵעַל הָאָרֶץ וַיָּסַר נֹחַ אֶת־מִכְסֵה הַתֵּבָה וַיַּרְא וְהִנֵּה חָרְבוּ פְּנֵי הָאֲדָמָה׃

ḥå̆rəḇ-ū *ham=mayim* *mē=ʿal* *hå̆='å̆rɛṣ*
dry_up\PFV-3PL the=water(PL) from=on the=earth

way-y-å̆sar-Ø *nōaḥ* *'ɛṯ* *miḵsɛ*
WAYQ-3M-move_away\CAUS-SG Noah OBJ covering.of

hat=tēḇå̆ *way-y-ar'-Ø* ***wə=hinnē*** *ḥå̆rəḇ-ū* *pən-ē*
the=ark WAYQ-3M-see-SG **and=behold** dry_up\PFV-3PL face(PL)-of

hå̆='ă̆ḏå̆måå̆
the=earth

'The waters dried up from the earth. Then Noah removed the covering of the ark, and saw — *and behold!* — the surface of the earth was drying.'

I consider cases of this latter type identical: the Addressee (the reader) takes the perspective of the subject of *rå̄'å̄* 'see', as it were, and thus sees

17 Also Gen. 19:28; Jdg. 21:21. With a participle the complement can be either clausal or nominal: Gen. 18:2; 24:63; 26:8; 29:2; 37:25; Jdg. 3:24; 9:43; 1 Sam. 10:11; 2 Sam. 13:34. The following are ambiguous between participle and suffix conjugation: Gen. 33:1; 2 Sam. 18:24.

18 Alternatively, based on manuscript evidence, emend אַחַר to אֶחָד *'ɛḥå̆ḏ* 'one'.

19 The distinction is not always clear. In Gen. 19:28, Abraham sees the smoke of Sodom and Gomorrah, of which the existence can be deduced by the Addressee from the sulfur and fire in 19:24. The smoke itself has, however, not been mentioned yet. All instances are mentioned in footnote 17, whether the information is only new to the subject or also to the Addressee.

the information as new. In this way the writer involves the Addressee (the reader) in the story.

An event being seen can also be described with a nominal complement, when modified by a participle used attributively:

(4.8) 2 Sam. 11:2: וַיַּ֣רְא אִשָּׁ֔ה רֹחֶ֖צֶת מֵעַ֣ל הַגָּ֑ג

way-y-ar'-Ø *'iššå* *rōḥɛṣ-ɛṯ* *mē='al* *hag=gå̄g*
WAYQ-3M-see-SG woman bathe\PTCP-F.SG from=on the=roof

'… and he (David) saw a woman bathing from upon the roof.'

This strategy can be combined with *hinnē* 'behold', in which case the information is noteworthy (e.g. Gen. 18:2). When *hinnē* 'behold' is not used (as in [4.8]), the information is new and unexpected (hence not easily accommodated), but also not marked as particularly noteworthy or immediately requiring the Addressee's attention; this strategy thus provides a middle ground between complementation with *kī* on the one hand, and *hinnē* on the other.

In conclusion, the division of labor is clear: interrogatives mark a lack of information, *hinnē* 'behold' introduces discourse-new or noteworthy information, and *kī* introduces discourse-old or easily accommodated information. A nominal complement modified by an attributively used participle, unless combined with *hinnē* 'behold', is used for information that is new and not expected, but also not particularly noteworthy. This is schematized in (4.9).

(4.9)

kī object clause	Nominal complement modified by participle	*wə-hinnē* 'and behold' + clause
← Information is in Common Ground		Information is new and unexpected or noteworthy →

4.3.2 Cognitive verbs

The class of cognitive verbs consists primarily of יָדַע *yå̄ḏaʿ* 'know' (72 times); other verbs in this class are נִחַם *niḥam* 'regret' (4 times), בִּין *bīn* 'understand' (twice), כִּחֵד *kiḥēḏ* 'hide', and זָכַר *zå̄ḵar* 'remember' (both once). All these less frequent predicates obviously refer to discourse-old information in the Common Ground, for example:

(4.10) 1 Sam. 15:35b: נִחַמְתִּי כִּי־הִמְלַכְתִּי אֶת־שָׁאוּל לְמֶלֶךְ

niḥam-tī ***kī*** *himlaḵ-tī* *ʾɛṯ* *šɑ̊ʾūl* *lə=mɛlɛḵ*
regret\PFV-1SG **COMP** be_king\CAUS.PFV-1SG OBJ Saul to=king

'I have come to regret *that* I have made Saul king.'

The verb *yɑ̊ḏaʿ* 'know' is often used in the same sense as *rɑ̊ʾɑ̊* 'see', meaning 'realize', and then has a complement that is obviously discourse-old or accommodated by the Addressee. However, *yɑ̊ḏaʿ* 'know' can be used more easily to impose information on the Common Ground:

(4.11) 1 Sam. 28:1: וַיֹּאמֶר אָכִישׁ אֶל־דָּוִד יָדֹעַ תֵּדַע כִּי אִתִּי תֵּצֵא בַּמַּחֲנֶה

way-y-ōʾmɛr-Ø *ʾɑ̊ḵīš* *ʾɛl* *dɑ̊wiḏ* *yɑ̊ḏōaʿ* *t-ēḏaʿ-Ø*
WAYQ-3M-say-SG Achish to David know\INFABS 2-know\IPFV-M.SG

kī *ʾitt-ī* *t-ēṣēʾ-Ø* *b=am=maḥănɛ*
COMP with-me 2-go_out\IPFV-M.SG in=the=camp

'And Achish said to David: "You should know for sure *that* you will go out with me in battle."'

Here, David has sought refuge with the Philistine king Gath. When the Philistines prepare to fight Israel, Achish makes sure that David understands that he has to fight on Achish's side now, and cannot refuse to participate. This is not something they discussed before. Using *kī* here imposes the information on the Common Ground, which works well for a command, and is reinforced by the infinitive absolute. Other than cases of imposition, however, *kī*-clauses with cognitive verbs are quite similar to those with verbs of perception.

4.3.3 Speech verbs

The most frequent speech verb with *kī*-clauses is הִגִּיד *higgīḏ* 'inform' (13 times, and 5 times in the passive voice); we also have נִשְׁבַּע *nišbaʿ* 'swear' (2 times), and אָמַר *ʾɑ̊mar* 'say', בִּשֵּׂר *biśśēr* 'bring good news', and גָּלָה אֹזֶן *gɑ̊lɑ̊ ʾōzɛn* 'uncover the ear' (1 time each).[20] The large number of occurrences with

20 I follow Miller (2003: 103–116) in rejecting the idea that *kī* can introduce direct speech: *kī* either introduces indirect speech, or it is the first word in the reported direct speech. I discuss cases of the former here; when *kī* is part of the reported speech it can have any other function, which I discuss throughout the rest of this chapter.

higgīḏ 'inform' compared to the overall more common *ʾāmar* 'say' already reflects the fact that *kī* refers to the Common Ground, since the complement of *higgīḏ* 'inform' is more often already known to the Addressee:[21]

(4.12) Jdg. 9:46–47: וַיִּשְׁמְעוּ כָּל־בַּעֲלֵי מִגְדַּל־שְׁכֶם וַיָּבֹאוּ אֶל־צְרִיַח בֵּית אֵל בְּרִית׃ וַיֻּגַּד לַאֲבִימֶלֶךְ כִּי הֻתְקַבְּצוּ כָּל־בַּעֲלֵי מִגְדַּל־שְׁכֶם׃

way-y-išmaʿ-ū *kāl* *baʿal-ē* *mīḡəḏal* *šəḵɛm* *way-y-āḇōʾ-ū*
WAYQ-3M-hear-PL all.of lord-PL.of tower.of Shechem WAYQ-3M-come-PL

ʾɛl *ṣərīaḥ* *bēṯ* *ʾēl* *bərīṯ* *way-y-uggaḏ-Ø*
to stronghold.of house.of El Berith WAYQ-3M-inform\PASS-SG

la=ʾăḇīmɛlɛḵ ***kī*** *hīṯə-qabbəṣ-ū* *kāl* *baʿăl-ē*
to=Abimelech **COMP** REFL-gather\PLURACT.PFV-3PL all.of lord-PL.of

mīḡəḏal *šəḵɛm*
tower.of Shechem

'The leaders of the Tower of Shechem heard [it] and came to the stronghold of the house of El-Berith. And Abimelech was informed *that* all the leaders of the tower of Shechem had gathered.'

Indirect speech can also be introduced with the relative pronoun אֲשֶׁר *ʾăšɛr* (Miller 2003: 97–98), but since this is rare and mostly a feature of Late Biblical Hebrew, I will not compare the two here. A comparison with asyndetic indirect speech (Miller 2003: 119–123) is interesting, though. It should be noted that asyndetic indirect speech only occurs embedded within direct speech (Miller 2003: 120), and the absence of a complementizer may therefore be attributed in part to register (cf. the more frequent omission of the English complementizer *that* in direct speech: Elsness 1984; Rissanen 1991). However, when we compare indirect speech with and without complementizer, both embedded in direct speech, a Common Ground effect can be observed:

(4.13) Gen. 12:13: אִמְרִי־נָא אֲחֹתִי אָתְּ

ʾimr-ī *nāʾ* *ʾăḥōṯ-ī* *ʾāttə*
say\IMP-F.SG please sister-mine you

'(Abram said to his wife Sarai: "…) Please say Ø you are my sister."'

21 Keep in mind here, again, that the Addressee, for our purposes, is not the Addressee of the speech event (the one who is being informed), but the Addressee of the reporting of the speech event (typically, the reader).

(4.14) Gen. 12:18: וַיִּקְרָ֣א פַרְעֹ֔ה לְאַבְרָ֖ם וַיֹּ֣אמֶר ... לָמָ֣ה לֹא־הִגַּ֣דְתָּ לִ֔י כִּ֥י אִשְׁתְּךָ֖ הִֽוא׃

way-y-iqrå'-Ø *pār'ōh* *lə='aḇrå̄m* *way-y-ō'mɛr-Ø* ... *lå̄mmå̄* *lō'*
WAYQ-3M-call-SG Pharaoh to=Abram WAYQ-3M-say-SG ... why not

higgaḏ-tå̄ *l-ī* ***kī*** *'išt-əḵå̄* *hī'*
inform\PFV-2M.SG to-me **COMP** wife-yours she

'So Pharaoh called Abram and said: "... Why have you not told me *that* she is your wife?"'

In (4.13), Abram's idea to pretend that Sarai is his sister is new to the Addressee (Sarai), and there is no reason why it should be easily accommodated. As a result, *kī* is not used. However, in (4.14), Pharaoh has found out that Sarai is Abram's sister. The information is in his and Abram's Common Ground, and the indirect speech is introduced with *kī* accordingly.[22]

One could object that the pair in (4.13–4.14) is not minimal because two different verbs are used: *higgīḏ* 'inform' with *kī* in (4.14) and *'å̄mar* 'say' without in (4.13). Since asyndetic indirect speech complements are only ever found with *'å̄mar* 'say' (Miller 2003: 121), a better minimal pair cannot be given.[23] However, note that the simple fact that *kī* appears with one verb and not the other can and should be seen as reflecting the function of the complementizer to mark Common Ground: meaning 'inform', *higgīḏ* is simply much more suitable to talk about common knowledge, whereas *'å̄mar* 'say' is more often used in contexts where the Addressee is given new information, such as imperatives (4.13). This distribution therefore confirms the hypothesis concerning *kī*.

4.3.4 Miscellaneous complementation structures

Although nominalized *kī*-clauses are usually object clauses complementing verbs, they can also complement nouns and prepositions, or function as subject clauses. With nouns we only find עֵד *'ēḏ* 'witness' (1 Sam. 12:5; Ruth 4:9)

22 In Gen. 12:19, Pharaoh continues to ask: 'why did you say, "she is my sister"?', without *kī*. But *kī* is here excluded since the complement is a direct speech report, as can be seen from the pronominal reference (Miller 2003: 120).

23 The one example of *'å̄mar* 'say' with a *kī*-clause has the verb in the meaning 'think, say to oneself': 'I thought that you really hated [your bride], so I gave her to your best man' (Jdg. 15:2). Common Ground is imposed here; the Speaker informs the Addressee of an assumption he made based on the Addressee's earlier behavior.

and עֵדָה *ʿēḏǻ* 'legal proof' (Gen. 21:30b); with prepositions we only have עַד *ʿaḏ* 'until' (3 times).

The nouns *ʿēḏ* 'witness' and *ʿēḏǻ* 'legal proof' are used with *kī* only to establish who can corroborate a certain fact. This is always a known fact, that is, a fact in the Common Ground:

(4.15) 1 Sam. 12:4–5: וַיֹּאמְרוּ֙ לֹ֣א עֲשַׁקְתָּ֔נוּ וְלֹ֥א רַצּוֹתָ֖נוּ וְלֹא־לָקַ֣חְתָּ מִיַּד־אִ֑ישׁ מְאוּמָֽה׃
וַיֹּ֥אמֶר אֲלֵיהֶ֖ם עֵ֣ד יְהוָ֑ה בָּכֶ֗ם וְעֵד֙ מְשִׁיח֣וֹ הַיּ֔וֹם הַזֶּ֖ה כִּ֥י לֹ֛א מְצָאתֶ֥ם בְּיָדִ֖י מְאוּמָ֑ה

way-y-ōʾmər-ū *lōʾ* *ʿăšaq-tǻ-nū* *wə=lōʾ*
WAYQ-3M-say-PL not wrong\PFV-2M.SG-us and=not

raṣṣō-ṯǻ-nū *wə=lōʾ* *lǻqaḥ-tǻ* *miy=yaḏ*
oppress\PFV-2M.SG-us and=not take\PFV-2M.SG from=hand.of

ʾīš *məʾūmǻ* *way-y-ōʾmɛr-Ø* *ʾălē-hɛm* *ʿēḏ* *yhwh*
anyone anything WAYQ-3M-say-SG to-them witness Yahweh

bǻ-ḵɛm *wə=ʿēḏ* *məšîḥ-ō* *hay=yōm* *haz=zɛ* ***kī***
against-you and=witness annointed-his the=day the=this **COMP**

lōʾ *məṣǻʾ-ṯɛm* *bə=yǻḏ-ī* *məʾūmǻ*
not find\PFV-2M.PL in=hand-mine anything

'And they said: "You have not wronged us or oppressed us, and you have not taken anything from anyone." So he said to them: "Today Yahweh is a witness against you, and his anointed is a witness, *that* you have not found any [charge] against me."'

Here, the complement clause simply reiterates what has already been said in the previous verse. Without *kī*, *ʿēḏ* 'witness' can be used with new information. In the following example, it is clear that Laban and Jacob are going to form some kind of covenant, but the preceding context provides no information based on which the Addressee can deduce that the cairn will be a marker of the border between them. Therefore *kī* would be inappropriate:

(4.16) Gen. 31:52: עֵד הַגַּל הַזֶּה וְעֵדָה הַמַּצֵּבָה אִם־אָנִי לֹא־אֶעֱבָר אֵלֶיךָ אֶת־הַגַּל
הַזֶּה וְאִם־אַתָּה לֹא־תַעֲבֹר אֵלַי אֶת־הַגַּל הַזֶּה וְאֶת־הַמַּצֵּבָה הַזֹּאת לְרָעָה׃

ʿēḏ *hag=gal* *haz=zɛ* *wə=ʿēḏǻ* *ham=maṣṣēḇǻ* *ʾim* *ʾǻnī* *lōʾ*
witness the=cairn the=this and=witness the=pillar if I not

ʾ-ɛʿĕḇōr *ʾēlɛ-ḵǻ* *ʾɛṯ* *hag=gal* *haz=zɛ* *wə=ʾim* *ʾattǻ*
1SG-pass_over\IPFV to-you OBJ the=cairn the=this and=if you

lō’ ṯ-aʿăḇōr-Ø ’ēla-y ’eṯ hag=gal haz=zɛ wə’=ɛṯ
not 2-pass_over\IPFV-M.SG to-me OBJ the=cairn the=this and=OBJ
ham=maṣṣēḇå haz=zō’ṯ lə=råʿå
the=pillar the=this for=evil

(Laban to Jacob:) "'This cairn is a witness, and the pillar is a witness: Ø I will not pass this cairn to you, and you will not pass this cairn and this pillar to me to do harm.'"

Each instance with *ʿaḏ* 'until' (Gen. 26:13; 41:49; 2 Sam. 23:10) describes the direct consequence of the matrix clause. For instance, in (4.17), being very wealthy is a direct consequence of becoming more and more wealthy. The *kī*-clause is therefore easily accommodated in the Common Ground:

(4.17) Gen. 26:13: וַיִּגְדַּל הָאִישׁ וַיֵּלֶךְ הָלוֹךְ וְגָדֵל עַד כִּי־גָדַל מְאֹד׃

way-y-iḡdal-Ø hå=’īš way-y-ēlɛḵ-Ø hålōḵ
WAYQ-3M-be_great-SG the=man WAYQ-3M-go-SG go\INFABS
*wə=ḡåḏēl-Ø ʿaḏ **kī** ḡåḏal-Ø mə’ōḏ*
and=be_great\PTCP-M.SG until **COMP** be_great\PFV-3M.SG very

'The man was wealthy, and he became more and more wealthy[24] to the point *that* he was very wealthy.'

The remaining cases are subject clauses. Two of these begin with אַף כִּי *’ap̄ kī* '[it is] even [the case] that' (Gen. 3:1; 1 Sam. 14:30). I first discuss the other four, which are more straightforward: they all nominalize a previously introduced proposition, and thus refer to Common Ground (1 Sam. 25:30; 2 Sam. 9:1; 18:3; Ruth 2:22). For instance, in (4.18), the nominalized clause refers to the same proposition as the earlier 'you will not go out':

(4.18) 2 Sam. 18:2–3c: וַיֹּאמֶר הַמֶּלֶךְ אֶל־הָעָם יָצֹא אֵצֵא גַּם־אֲנִי עִמָּכֶם: וַיֹּאמֶר הָעָם לֹא תֵצֵא ... וְעַתָּה טוֹב כִּי־תִהְיֶה־לָּנוּ מֵעִיר לַעְזוֹר[25]׃

way-y-ōʾmɛr-Ø ham=mɛlɛḵ ’ɛl hå=ʿåm yåṣō’
WAYQ-3M-say-SG the=king to the=people go_out\INFABS
’-ēṣē’ gam ’ănī ʿimm-å̄ḵɛm way-y-ōʾmɛr-Ø
1SG-go_out\IPFV also I with-you WAYQ-3M-say-SG

24 Reading *ḡåḏōl* (INFABS) for *ḡåḏēl*; for the durative interpretation cf. Gzella (2008).

25 The consonantal text has לעזיר; a misspelling or the same form with distant assimilation (Tsumura 2014: 137–138).

hă̄=ʿām *lō'* *ṯ-ēṣē'-Ø* ... *wə=ʿattă̄* *ṭōḇ*
the=people(M) not 2-go_out\IPFV-M.SG ... and=now good
kī *ṯ-īhəyɛ-Ø* *lă̄-nū* *mē=ʿīr* *la=ʿzōr*
COMP 2-be\IPFV-M.SG for-us from=city for=help

'The king said to the people: "I myself will also go out with you." But the people said: "[You will not go out]$_i$ ...; now, it is better *that* [you provide support for us from the city]$_i$."'

These cases thus confirm the hypothesis that *kī* marks the use of Common Ground (unfortunately, however, they cannot be compared to asyndetic finite subject clauses; these do not exist).

The cases with *'ap̄ kī* '[it is] even [the case] that ...' are as follows:[26]

(4.19) Gen. 3:1: וַיֹּ֤אמֶר אֶל־הָֽאִשָּׁה֙ אַף֙ כִּי־אָמַ֣ר אֱלֹהִ֔ים לֹ֥א תֹאכְל֖וּ מִכֹּ֣ל עֵ֥ץ הַגָּֽן׃

way-y-ōʾmɛr-Ø *ʾɛl* *hă̄=ʾiššă̄* ***ʾap̄*** ***kī*** *ʾă̄mar-Ø* *ʾĕlōhīm*
WAYQ-3M-say-SG to the=woman **even** **COMP** say\PFV-3M.SG God
lō' *ṯ-ōʾḵəl-ū* *mik=kōl* *ʿēṣ* *hag=gă̄n*
not 2-eat\IPFV-M.PL from=all.of tree.of the=garden

'And [the snake] said to the woman: "Is it *really* the case *that* God has said: 'You shall not eat from any tree of the garden'?"'

(4.20) 1 Sam. 14:30a: אַף֩ כִּ֨י לוּא֙ אָכַ֣ל הַיּ֔וֹם הָעָ֖ם מִשְּׁלַ֣ל אֹיְבָ֑יו אֲשֶׁ֥ר מָצְא֖וּ כִּ֥י עַתָּ֛ה לֹא־רָבְתָ֥ה מַכָּ֖ה בַּפְּלִשְׁתִּֽים׃

ʾap̄ ***kī*** *lū'* *ʾă̄ḵōl* *ʾă̄ḵal-Ø* *hayyōm* *hă̄=ʿām*
even **COMP** had_not eat\INFABS eat\PFV-3M.SG today the=army(M)
miš=šəlal *ʾōyəḇ-ā̆yw* *ʾăšɛr* *mă̄ṣă̄'-Ø* *kī* *ʿattă̄*
from=provision.of enemy-PL.its REL find\PFV-3M.SG COMP now
lō' *ră̄ḇə-ṯă̄* *makkă̄* *b=ap=pəlištī-m*
not be_great\PFV-3F.SG slaughter(F) against=the=Philistine-PL

'(See how my eyes gleamed when I tasted just a little of this honey.) It's *certainly* the case *that*, had the army today eaten from the enemies' provision which it found, that now the slaughter of the Philistines would have been greater.'

Example (4.19) is a case of imposed Common Ground: the snake presumably knows that God has only forbidden the people to eat from one tree,

26 The construction *'ap̄ kī* can sometimes be read as 'how much more/less', introducing a clause that provides a stronger reason for an implicit assertion than the reason provided in the previous clause (Van der Merwe et al. 2017: §40.14.1b). Three cases where *'ap̄ kī* should be read together are classified as causal (1 Sam. 21:6b; 23:3; 2 Sam. 16:11a).

but gains the woman's trust by pretending he is asking a simple question. By pretending that he is ill-informed, the snake presents itself as harmless to the woman, which it will subsequently exploit.

In (4.20), Common Ground is accommodated. In the previous clause, the Speaker has suggested that he was strengthened by eating just a little of the enemies' provisions; the following verse simply extends this to the rest of the army.

4.3.5 Summary

In conclusion, both the distribution of complementizer *kī* over different matrix predicates and a comparison with other complementation strategies support the hypothesis that *kī* marks information in the Common Ground, or information that is easily accommodated in the Common Ground. In terms of distribution we may note the frequent use with רָאָה *rāʾā* in the sense 'realize (old information)' rather than the literal 'see (something new)', as well as the preference for הִגִּיד *higgīḏ* 'inform (of old information)' over אָמַר *ʾāmar* 'say (something new)'. I have compared *kī* to various other complementation strategies, such as וְהִנֵּה *wə-hinnē* 'and behold' and asyndetic indirect speech, which can all be shown to be used when the complement is not in the Common Ground yet, in contrast to the cases with *kī*. The following sections proceed with the analysis of adverbial *kī*-clauses.

4.4 Causal *kī*

As mentioned in section 4.2, *kī* most frequently introduces a causal clause, which gives the cause, reason, or ground for the event described in the main clause.[27] In the majority of cases (almost 90%), the cause given in the clause is either already in the Common Ground or easily accommodated by the Addressee.[28] It is not uncommon for causal conjunctions to be sensitive to

27 The discussion of causal-adversative *kī* is delayed until section 4.5, where adversative *kī* is discussed as well.

28 Some brief remarks are in order about my classification of causal *kī* when a proper name is explained. There are 23 cases in my corpus, of which 20 in Genesis. In the data set these are marked as "naming": Gen. 2:23; 3:20; 4:25a; 10:25; 11:9; 16:11, 13; 17:5, 15; 21:31; 26:20, 22; 29:32ab, 33a; 32:29b, 31; 35:7; 41:51, 52; 1 Sam. 1:20; 2 Sam. 7:27; Ruth 1:20. One of these is causal-adversative with no reference to Common Ground (Gen. 17:15). The other 22 are causal: 17 with reference to discourse-old information (Gen. 2:23; 3:20;

reference to Common Ground. As mentioned in passing in the introduction, similar observations apply to, for example, English *since* (e.g. Dancygier & Sweetser 2000). Examples of such straightforward cases of causal *kī* are given in (4.21–4.24). In (4.21), the reason is in the Common Ground, because it is cultural knowledge which the Speaker (writer) assumes the Addressee (reader) to have. In (4.22), the reason is in the Common Ground because it has been given in the previous clause.

(4.21) Gen. 34:14: לֹ֣א נוּכַ֗ל לַעֲשׂוֹת֙ הַדָּבָ֣ר הַזֶּ֔ה לָתֵ֥ת אֶת־אֲחֹתֵ֖נוּ לְאִ֣ישׁ אֲשֶׁר־לֹ֣ו עָרְלָ֑ה כִּ֥י־חֶרְפָּ֖ה הִ֥וא לָֽנוּ׃

lō' *n-ūḵal* *la=ʿăśōṯ* *had=dāḇār* *haz=zɛ* *lā̆=ṯēṯ* *ʾeṯ*
not 1PL-be_able\IPFV to=do\INF the=thing the=this to=give\INF OBJ

ʾăḥōṯ-ēnū *lə=ʾīš* *ʾăšɛr* *l-ō* *ʿărlā̆* ***kī*** *ḥɛrpā̆*
sister-ours to=man REL for-him uncircumcised **COMP** disgrace

hī' *lā̆-nū*
it for-us

'We cannot do this, giving our sister to a man who is uncircumcised, *for* it is a disgrace to us.'

(4.22) Jdg. 14:17: וַתֵּ֤בְךְּ עָלָיו֙ שִׁבְעַ֣ת הַיָּמִ֔ים אֲשֶׁר־הָיָ֥ה לָהֶ֖ם הַמִּשְׁתֶּ֑ה וַיְהִי֙ | בַּיֹּ֣ום הַשְּׁבִיעִ֔י וַיַּגֶּד־לָ֖הּ כִּ֣י הֵצִ֣יקָה אֹתֹ֑ו

wat-t-ēḇk-Ø *ʿā̆l-ā̆yw* *šiḇʿa-ṯ* *hay=yā̆m-īm* *ʾăšɛr* *hā̆yā̆-Ø*
WAYQ-F-cry-3SG on-him seven-of the=day-PL REL be\PFV-3M.SG

lā̆-hɛm *ham=mištɛ* *wa-y-əhī-Ø* *b=ay=yōm* *haš=šəḇīʿī*
for-them the=party(M) WAYQ-3M-be-SG on=the=day the=seventh

way-y-aggɛḏ-Ø *l-ā̆h* ***kī*** *hĕṣīq-aṯ-hū*
WAYQ-3M-inform-SG to-her **COMP** press\PFV-3F.SG-him

'She cried in his presence during the seven days of their party, and on the seventh day he told her, *because* she pressed him so.'

4:25a; 11:9; 16:13; 17:5; 21:31; 26:20, 22; 29:32ab, 33a; 32:29b, 31; 35:7; 1 Sam. 1:20; 2 Sam. 7:27); 4 with accommodated Common Ground (Gen. 16:11; 41:51, 52; Ruth 1:20); and one without reference to Common Ground (Gen. 10:25). What sets the cases with *kī* apart from other constructions to give the reason for a proper name is that there is clear wordplay involved when *kī* is used. For example (Gen. 16:13): '[Hagar] called Yahweh who spoke to her, "you are God-seeing-me (אֵל רֳאִי *ʾēl rŏʾī*)", *because* she said: "Have I not seen here after the one who sees me (רֹאִי *rōʾī*)?"'. What Hagar actually said (the direct speech in the *kī*-clause) can therefore be deduced more or less from the first clause, and should not be seen as new information. Without *kī*, there is no wordplay, or it is left unexplained (e.g. Gen. 25:25–26; 31:49; 35:18), so that the reason for the name cannot be accommodated based on the name itself.

In (4.23), it is said that the grain that Joseph is storing is 'a very great quantity', and eventually he has to stop counting it. Based on this, the information in the *kī*-clause (that the stored grain had become immeasurable) is easily accommodated.

(4.23) Gen. 41:49b: וַיִּצְבֹּר יוֹסֵף בָּר כְּחוֹל הַיָּם הַרְבֵּה מְאֹד עַד כִּי־חָדַל לִסְפֹּר כִּי־אֵין מִסְפָּר׃

way-y-iṣbōr-Ø	*yōsēp*	*bɑ̊r*	*kə=ḥōl*	*hay=yɑ̊m*
WAYQ-3M-store-SG	Joseph	grain	like=sand.of	the=sea

harbɛ	*məʾōḏ*	*ʿaḏ*	*kī*	*ḥɑ̊ḏal-Ø*	*li=spōr*
be_great\CAUS.INF	very	until	COMP	stop\PFV-3M.SG	to=count\INF

kī	*ʾēn*	*mispɑ̊r*
COMP	not_exist	number

'And Joseph stored grain — as much as the sand of the sea, a very great quantity — to the point that he stopped counting it *because* it was immeasurable.'

Similarly, in (4.24), Samuel has to take oil and go to Jesse. Based on this (and the fact that Saul has been rejected as a king in the preceding chapter), it is easily accommodated that someone in Jesse's family will be the new king.

(4.24) 1 Sam. 16:1: מַלֵּא קַרְנְךָ שֶׁמֶן וְלֵךְ אֶשְׁלָחֲךָ אֶל־יִשַׁי בֵּית־הַלַּחְמִי כִּי־רָאִיתִי בְּבָנָיו לִי מֶלֶךְ׃

mallēʾ-Ø	*qarn-əḵɑ̊*	*šɛmɛn*	*wə=lēḵ-Ø*
fill\IMP-M.SG	horn-yours	oil	and=go\IMP-M.SG

ʾ-ɛšəlɑ̊ḥ-ɑ̆ḵɑ̊	*ʾɛl*	*yišay*	*bēṯ hallaḥmī*
1SG-send_out\IPFV-you.OBJ	to	Jesse	the_Bethlehemite

kī	*rɑ̊ʾī-ṯī*	*bə=ḇɑ̊n-ɑ̊yw*	*l-ī*	*mɛlɛḵ*
COMP	see\PFV-1SG	in=son-PL.his	for-me	king

'Fill your horn with oil and go, I will send you to Jesse the Bethlehemite, *for* I have seen a king for me amongst his sons.'

Things become more interesting in cases where Common Ground is imposed. In the following example, David has just sneaked into Saul's camp while Abner was on guard. David then says to Abner that he could have killed Saul. Using *kī*, David pretends that Abner should know that an enemy soldier (David himself) came into the camp, thus emphasizing that Abner did not do a very good job protecting Saul:

(4.25) 1 Sam. 26:15: וְלָמָּה לֹא שָׁמַרְתָּ אֶל־אֲדֹנֶיךָ הַמֶּלֶךְ כִּי־בָא אַחַד הָעָם לְהַשְׁחֵית אֶת־הַמֶּלֶךְ אֲדֹנֶיךָ׃

wə=lå̄mmå̄ lō' šå̄mar-tå̄ 'ɛl 'ăḏōnɛ-ḵå̄ ham=mɛlɛḵ ***kī***
and=why not guard\PFV-2M.SG to lord-yours the=king **COMP**
ḇå̄'-Ø 'aḥaḏ hå̄=ʿå̄m lə=hašḥīṯ 'ɛṯ ham=mɛlɛḵ
come\PFV-3M.SG one.of the=army to=destroy\INF OBJ the=king
'ăḏōnɛ-ḵå̄
lord-yours

'And why haven't you protected your lord the king, *given that* a soldier came to kill the king your lord?'

In all the examples in (4.21–4.25), the Speaker positions a sentential complement close to the Addressee using [+distal] *kī*. Depending on the context, this can have several effects: it may signal to the Addressee that discourse-old information is being referenced (4.21–4.22), or the Addressee may be prompted to accommodate some new information in the Common Ground (4.23–4.24), or the Speaker may present information as something the Addressee should have known, knowing full well that they do not (4.25).

As further evidence for the fact that causal *kī* marks Common Ground, it is interesting to see that when a cause consists of partially new information, the new information can be introduced by *wə-hinnē* 'and behold'.[29] In (4.26), the theory predicts that simply וְ *wə* 'and' would be infelicitous, because that would incorrectly suggest that the Addressee already knows that the land is good:

(4.26) Jdg. 18:9: קוּמָה וְנַעֲלֶה עֲלֵיהֶם כִּי רָאִינוּ אֶת־הָאָרֶץ וְהִנֵּה טוֹבָה מְאֹד

qūm-å̄ *wə=n-aʿălɛ* *ʿălē-hɛm* ***kī*** *rå̄'ī-nū*
stand_up\IMP-M.SG and=1PL-go_up\IPFV to-them **COMP** see\PFV-1PL
'ɛṯ hå̄='å̄rɛṣ ***wə=hinnē*** *ṭōḇ-å̄* *mə'ōḏ*
OBJ the=land(F) **and=behold** good-F.SG very

'(And the Danites returned to their brothers in Zorah and Eshtaol, and their brothers said to them: "How did it go?" And they said:) "Come on, let's go up against them, *for* we saw their land, *and look*: it's very good!"'

29 McCarthy (1980: 333–334) claims that *wə-hinnē* 'and behold' can have a causal sense and practically replace *kī*, but the examples are not convincing so I will not compare the two.

4.4.1 Backgrounded causal clauses

There are, however, cases of causal *kī* where Common Ground is not being referred to or even imposed. Here the original [+distal] feature of *kī* is often still relevant. There are plenty of cases where the cause is backgrounded, and in that sense placed at a distance from the main topic of conversation. Quite often the *kī*-clause provides the reason for a positive or negative command (16 out of 39 cases without reference to Common Ground):[30]

(4.27) Jdg. 13:5b: הִנְּךָ הָרָה וְיֹלַדְתְּ בֵּן וּמוֹרָה לֹא־יַעֲלֶה עַל־רֹאשׁוֹ כִּי־נְזִיר אֱלֹהִים יִהְיֶה הַנַּעַר מִן־הַבָּטֶן

hinn-ǎ̱k *hǎr-ǎ* *wə-yōlaḏ-t* *bēn* *ū=mōrǎ* *lō’*
look-you pregnant-F.SG WQAT-bear-2F.SG son and=razor(M) not

y-aʿălɛ-Ø *ʿal* *rōʾš-ō* ***kī*** *nəzīr* *ʾĕlōhīm*
3M-go_up\IPFV-SG on head-his **COMP** dedicated.of God

y-ihyɛ-Ø *han=naʿar* *min* *hab=bǎ̱ṭɛn*
3M-be\IPFV-SG the=boy from the=womb

‘Look, you are pregnant and will give birth to a son, but a razor shall not come on his head, *for* the boy will be dedicated to God from the womb.’

(4.28) Gen. 43:16: הָבֵא אֶת־הָאֲנָשִׁים הַבֵּיְתָה וּטְבֹחַ טֶבַח וְהָכֵן כִּי אִתִּי יֹאכְלוּ הָאֲנָשִׁים בַּצָּהֳרָיִם׃

hǎ̱bē’-Ø *ʾɛṯ* *hǎ̱=ʾănǎ̱š-īm* *hab=bǎ̱yəṯ-ǎ̱*
bring\IMP-M.SG OBJ the=man-PL the=house-ALL

ū=ṭəḇōaḥ-Ø *ṭɛḇaḥ* *wə=hǎ̱ḵēn-Ø* ***kī***
and=slaughter\IMP-M.SG animal and=prepare\IMP-M.SG **COMP**

ʾitt-ī *y-ōʾḵəl-ū* *hǎ̱=ʾănǎ̱š-īm* *b=aṣ=ṣǎ̱hǎrǎ̱yim*
with-me 3M-eat\IPFV-PL the=man-PL at=the=noon

‘Bring the men to the house and slaughter an animal and prepare it, *for* the men will eat with me at noon.’

In such cases, the *kī*-clause provides an explanation, but it is most important to the Speaker that the command is followed. For this reason, the explanation in the causal clause can be backgrounded and hence marked

30 In the data set these are marked with “command”. Positive commands: Gen. 21:12; 31:12; 40:15a; 43:16; 1 Sam. 14:39a; 16:11; 23:27; Ruth 3:18a. Negative commands: Gen. 2:17; 21:17; 26:24; 35:17; Jdg. 13:5b; 1 Sam. 4:20; 16:7a; 2 Sam. 13:32a.

with [+distal] *kī*. In the remaining cases, the *kī*-clause often provides information that is not crucial for the main story line and can therefore be seen as backgrounded as well.[31]

4.4.2 Lexicalized causal meaning

However, there remain some exceptions:

(4.29) Gen. 25:21: וַיֶּעְתַּ֥ר יִצְחָ֛ק לַיהוָ֖ה לְנֹ֣כַח אִשְׁתּ֑וֹ כִּ֥י עֲקָרָ֖ה הִ֑וא וַיֵּעָ֤תֶר לוֹ֙ יְהוָ֔ה וַתַּ֖הַר רִבְקָ֥ה אִשְׁתּֽוֹ׃

way-y-ɛʿtar-Ø *yiṣḥå̄q* *l=yhwh* *lənōḵaḥ* *ʾišt-ō*
WAYQ-3M-pray-SG Isaac to=Yahweh on_behalf_of wife-his

kī *ʿå̆qå̄r-å̄* *hīʾ* *way-y-ēʿå̄ṯɛr-Ø* *l-ō* *yhwh*
COMP barren-F.SG she WAYQ-3M-listened-SG to-him Yahweh

wat-t-ahar-Ø *riḇqå̄* *ʾišt-ō*
WAYQ-F-conceive-3SG Rebekah wife-his

'And Isaac prayed to Yahweh on behalf of his wife, *since* she was barren, and God heard his prayer and Rebekah his wife conceived.'

In (4.29), it is not known to the Addressee that Rebekah is barren, nor is there any reason why it should be easily accommodated. However, the information is not backgrounded either, since it is picked up at the end of the verse: 'and Rebekah his wife conceived'. For the 10 exceptions of this type I have no explanation based on a synchronic interpretation of the [+distal] feature.[32] I propose that the causal meaning of *kī* is lexicalized on the basis of the examples where Common Ground or distancing is relevant, so that the causal meaning could then be extended to other contexts.

Note that the causal meaning is more than frequent enough for such lexicalization to have taken place. Furthermore, the lexicalization is plausible since there is a clear developmental path. In an earlier stage of the language, there were only causal instances that are derived from a [+distal] feature (such as examples [4.21–4.28]). Speakers then reanalyzed *kī* as a simple marker of causation. This allowed for the spread to cases where the [+distal] feature does not seem to be interpretable anymore (4.29).

31 In the data set these are marked with "backgrounded": Gen. 5:24; 10:25; 15:16; 21:16; 42:4; Jdg. 4:3; 16:17; 1 Sam. 20:26a; 30:12; 2 Sam. 13:2; 14:15; Ruth 1:6a; 3:17.

32 In the data set these are marked as "lexicalized meaning", for reasons explained below: Gen. 25:21; 37:17; Jdg. 6:30ab; 1 Sam. 4:13; 6:19a; 13:19; 30:6a; 2 Sam. 6:6; 19:27a.

A critic may argue that if we need to assume lexicalization to have occurred for some cases anyway, it is simpler to describe the causal meaning of *kī* as lexicalized in all instances, including when the [+distal] feature is interpretable. On such a view, the fact that the *kī*-clauses in (4.21–4.28) expresses information that is already in the Common Ground, or is easily accommodated/imposed, is accidental. This argument cannot be maintained. First of all, this alternative is not, in fact, simpler, since there is no reason why a causal interpretation should not be pragmatically inferred, just like other adverbial interpretations (see sections 4.5 and 4.6). More importantly, however, this alternative approach cannot account for the distribution of *kī*. Only 2% of the total number of causal *kī*-clauses is exceptional in this way, and in all other cases the cause is in the Common Ground, easily accommodated, or imposed. A quick look at a dictionary shows that this is a much larger portion of the data than is the case with other causal conjunctions (e.g. יַעַן *yaʿan*), and similarly it can be shown that *kī* is much more frequent with commands than other causal conjunctions. These distributional facts can only be explained if *kī* has not simply lexicalized a causal meaning, but if many of the causal instances of *kī* are in fact pragmatically inferred from the general meaning of referring to Common Ground.

4.5 Adversative and causal-adversative *kī*

An adversative clause provides a contrast. For example, in (4.30), the answer of the people contrasts with Samuel's advice not to appoint a king. Since the people already asked for a king in verse 5, this request is in the Common Ground and here marked by *kī*:

(4.30) 1 Sam. 8:19: וַיְמָאֲנוּ הָעָם לִשְׁמֹעַ בְּקוֹל שְׁמוּאֵל וַיֹּאמְרוּ לֹּא כִּי אִם־מֶלֶךְ יִהְיֶה עָלֵינוּ׃

wa-y-əmåʾăn-ū *hå=ʿåm* *li=šmōaʿ* *bə=qōl* *šəmūʾēl*
WAYQ-3M-refuse-PL the=people to=listen\INF in=voice.of Samuel

way-y-ōʾmər-ū *lō'* ***kī*** *ʾim* *mɛlɛ̱k* *y-ihəyɛ-Ø* *ʿålē-nū*
WAYQ-3M-say-PL no **COMP** if king 3M-be\IPFV-SG over-us

'But the people refused to listen to Samuel and said: "no; *but* a king shall rule (lit.: be) over us!"'

I classified cases where the contrasting clause provides the reason or cause for a preceding negative statement as causal-adversative. Here again, the adverbial clause is in the Common Ground, because its propositional content has already been introduced in the preceding verse:

(4.31) Gen. 45:8: וְעַתָּה לֹא־אַתֶּם שְׁלַחְתֶּם אֹתִי הֵנָּה כִּי הָאֱלֹהִים

wə=ʿattå lō' 'attɛm šəlaḥ-tɛm 'ōṯ-ī hēnnå ***kī*** *hå=ʾĕlōhīm*
and=now not you send\PFV-2M.PL OBJ-me here **COMP** the=God

'(God sent me ahead of you …) So now, it is not you who have sent me here, *but/because* [it is] God.'

A comparison with other adversative strategies again shows that the degree of reference to Common Ground is especially high when *kī* is used. When the adversative clause presents new information, we often find the more neutral conjunction וְ *wə* 'and, but, …' instead:[33]

(4.32) Gen. 2:16–17: מִכֹּל עֵץ־הַגָּן אָכֹל תֹּאכֵל: וּמֵעֵץ הַדַּעַת טוֹב וָרָע לֹא תֹאכַל מִמֶּנּוּ

mik=kōl ʿēṣ hag=gå̄n ʾå̄ḵōl t-ō'ḵēl-Ø
from=all.of tree.of the=garden eat\INFABS 2-eat\IPFV-M.SG

ū=*mē=ʿēṣ had=daʿaṯ ṭōḇ wå̄=rå̄ʿ lō'*
and=from=tree.of the=knowledge.of good and=evil not

ṯ-ō'ḵal-Ø mimmɛn-nū
2-eat\IPFV-M.SG from-it

'You may eat from all the trees of the garden, *but* from the tree of the knowledge of good and evil you may not eat.'

A more marked way to introduce an adversative clause with new information content is to use רַק *raq* 'however, but'.[34] In the following example, the clause introduced by *raq* provides information that is not yet in the Common Ground:

33 Simply searching for "but" in an English translation yields many more examples where *wə* introduces new information (e.g. Gen. 6:8, 18; 8:9).

34 However, אַךְ *ʾaḵ* 'indeed, just, only, but' regularly refers to discourse-old information content when used adversatively (see Levinsohn 2011: 92–94 for cases where adversative *ʾaḵ* is used to emphasize a previously introduced point). Though *ʾaḵ* has cognates in Tigrē and Geʿez, I am not aware of a widely accepted etymology. It is tempting to associate this particle, like *kī*, with **ka*, but this is not necessary, as the existence of some other particles sensitive to the contents of the Common Ground is not in itself problematic for my hypothesis.

(4.33) Exod. 8:24: אָנֹכִי אֲשַׁלַּח אֶתְכֶם וּזְבַחְתֶּם לַיהוָה אֱלֹהֵיכֶם בַּמִּדְבָּר רַק הַרְחֵק לֹא־תַרְחִיקוּ לָלֶכֶת

ʾănōḵī ʾ-ăšallaḥ ʾɛṯ-ḵɛm ū-zəḇaḥ-tɛm l=yhwh
I 1SG-send_away\IPFV OBJ-you WQAT-sacrifice-2M.PL to=Yahweh

*ʾĕlōhē-ḵɛm b=am=miḏbǎr **raq** harḥēq lō'*
God-yours in=the=desert **however** be_far\CAUS.INFABS not

ṯ-arḥīq-ū lǎ=lɛḵɛṯ
2-be_far\CAUS.IPFV-M.PL to=go\INF

'I will let you go so that you may sacrifice to Yahweh your God in the desert, *however*, be sure not to go very far.'

Though most instances of adversative and causal-adversative *kī* make reference to Common Ground, both categories are characterized by a relatively high number of instances that cannot be explained using a [+distal] feature. In both adversative (4.34) and causal-adversative (4.35), the *kī*-clause provides information that is not easily accommodated and cannot be seen as backgrounded either:

(4.34) Jdg. 1:19a: וַיֹּרֶשׁ אֶת־הָהָר כִּי לֹא לְהוֹרִישׁ אֶת־יֹשְׁבֵי הָעֵמֶק כִּי־רֶכֶב בַּרְזֶל לָהֶם׃

*way-y-ōrɛš-Ø ʾɛṯ hǎ=hǎr **kī** lō' lə=hōrīš ʾɛṯ*
WAYQ-3M-conquer-SG OBJ the=hill **COMP** not to=conquer\INF OBJ

yōšəḇ-ē hǎ=ʿēmɛq kī rɛḵɛḇ barzɛl lǎ-hɛm
inhabit\PTCP-M.PL.of the=plain COMP chariot iron to-them

'And [Judah] conquered the hill country, *but* they could not conquer the people living in the plains, because they had iron chariots.'

(4.35) Gen. 24:3–4: לֹא־תִקַּח אִשָּׁה לִבְנִי מִבְּנוֹת הַכְּנַעֲנִי אֲשֶׁר אָנֹכִי יוֹשֵׁב בְּקִרְבּוֹ׃ כִּי אֶל־אַרְצִי וְאֶל־מוֹלַדְתִּי תֵּלֵךְ וְלָקַחְתָּ אִשָּׁה לִבְנִי לְיִצְחָק׃

lō' ṯ-iqqaḥ-Ø ʾiššǎ li=ḇn-ī mib=bən-ōṯ
not 2-take\IPFV-M.SG woman for=son-mine from=child-F.PL

*hak=kəna'ănī ʾăšɛr ʾǎnōḵī yōšēḇ-Ø bə=qirb-ō **kī***
the=Canaanite REL I live\PTCP-M.SG in=midst-its **COMP**

ʾɛl ʾarṣ-ī wə=ʾɛl mōlaḏt-ī ṯ-ēlēḵ-Ø
to land-mine and=to motherland-mine 2-go\IPFV-M.SG

wə-lǎqaḥ-tǎ ʾiššǎ li=ḇn-ī lə=yiṣḥǎq
WQAT-take-2M.SG woman for=son-mine for=Isaac

'You must not acquire a wife for my son from the women of the Canaanites among whom I am living, *but/because* you must go to my country and my motherland (instead) and take a wife for my son, for Isaac.'

These cases are therefore exceptional in the same way as causal (4.29) above. As with causal *kī*, a diachronic account remains necessary for these cases.

Given the relatively low number of occurrences of both adversative and causal-adversative, it is unlikely that these functions developed from the general meaning of *kī* directly. It is more likely that they are the result of a semantic shift, starting out with the already lexicalized causal meaning. It is easy to see how a causal meaning could expand to causal-adversative, and eventually lose the necessarily causal interpretation to become plain adversative; such a development is also widely attested cross-linguistically (cf. Locatell 2017: 247–248 and references therein).

In my corpus, lexicalized meanings are needed for 10 exceptional cases of causal *kī*, 4 cases of causal-adversative *kī*, and 5 cases of adversative *kī*.[35] Since the number of causal-adversative cases is roughly half of the number of causal cases, I consider it more likely that the lexicalized causal-adversative meaning (and subsequently, the adversative meaning) developed from the lexicalized causal meaning, than that it developed independently from the general meaning of *kī*.

4.6 Remaning adverbial uses

Having discussed causal and adversative *kī*-clauses in sections 4.4 and 4.5, respectively, this section reviews the evidence of the remaining, lower-frequency types of adverbial clauses introduced by *kī*: temporal and conditional clauses (section 4.6.1), resultative clauses (section 4.6.2), and concessive clauses (section 4.6.3).

In all these types of adverbial clauses, the information provided in the *kī*-clause is already in the Common Ground or easily accommodated, or (rarely) imposed on the Common Ground. There will therefore be no need to assume that any of these uses of *kī* are lexicalized, as expected given the low frequency of these categories. All instances can be derived synchronically from a [+distal] feature.

35 These are marked with "lexicalized meaning" in the data set. For the causal cases, see footnote 32. Causal-adversative cases are Gen. 17:15; 24:4; 32:29a; 35:10; adversative cases are Gen. 40:14; Jdg. 1:19a; 4:9a; 1 Sam. 15:35a; 2 Sam. 17:11.

4.6.1 Temporal and conditional clauses

In my corpus only few instances of *kī* are temporal, and even less conditional.[36] The distinction is not always clear-cut, as seen in (4.36). For this reason, I will discuss them together.

(4.36) Gen. 32:18: וַיְצַו אֶת־הָרִאשׁוֹן לֵאמֹר כִּי יִפְגָּשְׁךָ עֵשָׂו אָחִי וּשְׁאֵלְךָ לֵאמֹר

wa-y-əṣaw-Ø	*ʾɛṯ*	*hå̄=riʾšōn*	*lē=ʾmōr*
WAYQ-3M-command-SG	OBJ	the=first	to=say\INF

kī	*y-ip̄əgåš-Ø-ḵå̄*	*ʿēśå̄w*	*ʾå̄ḥ-ī*
COMP	3M-meet\IPFV-SG-you.OBJ	Esau	brother-mine

wi-šʾēl-Ø-əḵå̄	*lē=ʾmōr*
WQAT-ask-3M.SG-you.OBJ	to=say\INF

'He (Jacob) commanded the first [servant], saying, "*If/When* my brother Esau meets you and asks you, saying, '…'"'

There are two ways in which *kī* interacts with Common Ground in such cases. First, temporal and conditional clauses introduced by *kī* always describe an event that the Speaker considers certain or likely to occur or have occurred. *Kī* is not used for conditions that the Speaker does not expect to be met, or events that are unlikely to occur or have occurred. By modifying a main clause M with a conditional C, the Speaker can enlarge the Common Ground: M only applies given the current Common Ground augmented with the information that C is met. However, *kī* can only be used when this augmentation is reasonable given the current Common Ground: it should be easy to accommodate C (when it is likely or expected to occur or have occurred), or the Common Ground should already entail C (when it is certain to occur or have occurred). Second, in most occurrences, the event or condition described is also already salient in the surrounding discourse, and in that sense tracked by both Speaker and Addressee: it is under discussion, even if it is not certain that the condition is met.

Concretely, the first observation means that when *kī* introduces a conditional clause, it describes reasonably realistic conditions: it describes conditions that are likely to happen. In that sense, these conditional uses are much like the temporal uses of *kī*, and unlike the more general conditional use of אִם *ʾim* 'if' (cf. the similar contrast between English *when* and *if*):

36 This low frequency may be due to the narrative genre of my corpus (cf. Locatell 2017: 252 n. 322; most of his conditional examples are from legal texts, which I did not include).

(4.37) Gen. 18:26: אִם־אֶמְצָא בִסְדֹם חֲמִשִּׁים צַדִּיקִם בְּתוֹךְ הָעִיר וְנָשָׂאתִי לְכָל־הַמָּקוֹם בַּעֲבוּרָם׃

'im	*'-εmṣå'*	*bi=sḏōm*	*ḥămišš-īm*	*ṣaddīq-im*	*bə=ṯōḵ*
if	1SG-find\IPFV	in=Sodom	five-PL	righteous-PL	in=middle.of

hå=ʿīr	*wə-nåśå'-ṯī*	*lə=ḵål*	*ham=måqōm*	
the=city	WQAT-forgive-1SG	for=whole.of	the=place	

ba=ʿăḇūr-åm
for=sake-theirs

'*If/*When* I find at Sodom fifty righteous in the city, I will forgive the whole place for their sake.'

This is consistent with the fact that when *kī* and *'im* are combined, *kī* introduces a general condition, while *'im* introduces a more specific condition (Van der Merwe et al. 2017: §40.11(1a)). In such a construction, the condition in the *kī*-clause is more likely to be met than the condition in the *'im*-clause:

(4.38) Lev. 1:2–3: אָדָם כִּי־יַקְרִיב מִכֶּם קָרְבָּן לַיהוָה מִן־הַבְּהֵמָה מִן־הַבָּקָר וּמִן־הַצֹּאן תַּקְרִיבוּ אֶת־קָרְבַּנְכֶם׃ אִם־עֹלָה קָרְבָּנוֹ מִן־הַבָּקָר זָכָר תָּמִים יַקְרִיבֶנּוּ

'åḏåm	***kī***	*y-aqrīḇ-Ø*	*mikk-εm*	*qårbån*	
man	**COMP**	3M-present\IPFV-SG	from-you	offering	

l=yhwh	*min*	*hab=bəhēmå*	*min*	*hab=bå̊qå̊r*	*ū=min*
to=Yahweh	from	the=animals	from	the=herd	or=from

haṣ=ṣō'n	*t-aqrīḇ-ū*	*'εṯ*	*qårban-ḵεm*	***'im***
the=flock	2-present\IPFV-M.PL	OBJ	offering-yours	**if**

ʿōlå̊	*qårbå̊n-ō*	*min*	*hab=bå̊qå̊r*	*zå̊ḵå̊r*	*tå̊mīm*
burnt_offering	offering-his	from	the=herd	male	perfect

y-aqrīḇ-Ø-εnnū
3M-present\IPFV-SG-it.OBJ

'*If/When* (*kī*) a man amongst you brings a sacrifice to Yahweh, you must bring your offer from the animals of the herd or the flock. *If/*When* (*'im*) it is a burnt offering from the herd, he shall offer a male without blemish.'

This fits with the general meaning described for *kī* above: the adverbial clause does not introduce entirely unexpected information, but only information which is already expected, or at least easily accommodated given the Common Ground. In (4.36) above, with *kī*, Jacob's servant is going towards Esau and will therefore surely meet him, while in (4.37), with *'im*, it is not certain at all that there will be fifty righteous at Sodom. In (4.38), not all

sacrifices are burnt offerings, so the condition in the *ʾim*-clause is less likely to be met than the condition in the *kī*-clause.

Temporal clauses with future reference time can, in a way, be seen as an extreme case of such "expected conditionals"; they are essentially conditionals of which the condition is certain to be met at some future point in time:

(4.39) Gen. 31:49: יִצֶף יְהוָה בֵּינִי וּבֵינֶךָ כִּי נִסָּתֵר אִישׁ מֵרֵעֵהוּ׃

y-iṣɛp̄-Ø	*yhwh*	*bēn-ī*	*ū=b̲ēn-ɛk̲å̄*	***kī***
3M-guard\JUSS-SG	Yahweh	between-me	and=between-you	**COMP**

n-isså̄ṯ̲ēr	*ʾīš*	*mē=rēʿ-ēhū*
1PL-hide\MID.IPFV	one	from=companion-his

'May Yahweh watch between me and you *when* we are hidden from each other.'

Temporal clauses with past reference time are similar. They require the described event to have occurred; *kī* cannot be used, for example, for counterfactuals, which are typically marked by לוּ *lū* 'if only'.[37]

In some instances, a *kī*-clause with past reference time is frequentative ('whenever'), as in (4.40). This specific interpretation is only contextually inferred, however, and not contributed by *kī*, as it is not always available (4.41):[38]

(4.40) Jdg. 1:28: וַיְהִי כִּי־חָזַק יִשְׂרָאֵל וַיָּשֶׂם אֶת־הַכְּנַעֲנִי לָמַס וְהוֹרֵישׁ לְא הוֹרִישׁוֹ׃

37 Note that a counterfactual, like a temporal or conditional clause, can be described as augmenting the Common Ground. The only difference is that a counterfactual augments the Common Ground with information known to be false. If *kī* were simply described as *augmenting* the Common Ground, there would be no way to exclude a counterfactual interpretation. But since *kī* is [+distal] and therefore marks information that is in the Common Ground or at least easily accommodated, a counterfactual interpretation is excluded.

38 A frequentative interpretation is also possible in Jdg. 2:18a; 12:5; 16:16; 2 Sam. 6:13; and we have a durative interpretation ('while') in 1 Sam. 1:12; 17:48. Such interpretations are excluded in Gen. 6:1; 26:8; 31:37; 43:21; 44:24; Jdg. 6:5, 7; 8:1; 1 Sam. 14:29b; 22:22a; 2 Sam. 4:10, 11; 7:1, 12; 19:26. There does not appear to be a difference in the contribution of *kī* when preceded by וְהָיָה *wə-hāyā* 'and it will be' or וַיְהִי *wa-yəhī* 'and it was', compared to when *kī* stands alone. The contribution of these temporal markers can be seen as shifting the reference time (e.g. Van der Merwe et al. 2017: §40.24–25) independent from the discursive contribution made by *kī*.

wa-y-əhī-Ø ***kī*** *ḥăzaq-Ø* *yiśrāʾēl* *way-y-ăśɛm-Ø*
WAYQ-3M-be-SG **COMP** be_strong\PFV-3M.SG Israel WAYQ-3M-put-SG

ʾɛṯ *hak=kənaʿănī* *lă=mas* *wə=hōrēš* *lōʾ*
OBJ the=Canaanite to=forced_labor but=conquer\INFABS not

hōrīš-Ø-ō
conquer\PFV-3M.SG-it.OBJ

'And *whenever* Israel was strong they would put the Canaanites to work, but they did not totally conquer them.'

(4.41) Gen. 27:1: וַיְהִי֙ כִּֽי־זָקֵ֣ן יִצְחָ֔ק וַתִּכְהֶ֥יֶן עֵינָ֖יו מֵרְאֹ֑ת וַיִּקְרָ֞א אֶת־עֵשָׂ֧ו | בְּנ֛וֹ הַגָּדֹ֖ל

wa-y-əhī-Ø ***kī*** *zăqēn-Ø* *yiṣḥăq* *wat-t-iḵhɛ-nă*
WAYQ-3M-be-SG **COMP** be_old\PFV-3M.SG Isaac WAYQ-F-be_weak-3PL

ʿēn-ăyw *mē=rəʾōṯ* *way-y-iqrāʾ-Ø* *ʾɛṯ* *ʿēśāw* *bən-ō*
eye(F)-DU.his from=see\INF WAYQ-3M-call-SG OBJ Esau son-his

hag=gāḏōl
the=big

'And *when* Isaac was old and his eyes were too weak to see, he called Esau, his oldest son.'

To return to the second way in which temporal and conditional *kī* interacts with Common Ground: in most instances of temporal or conditional *kī*, the event or condition described in the adverbial clause is already salient in the surrounding discourse, and in that sense tracked by both Speaker and Addressee. Consider:

(4.42) Gen. 44:24: וַיְהִי֙ כִּ֣י עָלִ֔ינוּ אֶל־עַבְדְּךָ֖ אָבִ֑י

wa-y-əhī-Ø ***kī*** *ʿālī-nū* *ʾɛl* *ʿaḇd-əḵā* *ʾāḇ-ī*
WAYQ-3M-be-SG **COMP** go_up\PFV-1PL to servant-yours father-mine

'And *when* we went up to your servant my father, …'

(4.43) Jdg. 2:16–18: וַיָּ֤קֶם יְהוָה֙ שֹׁפְטִ֔ים … וְכִֽי־הֵקִ֥ים יְהוָ֛ה | לָהֶ֖ם שֹׁפְטִ֑ים

way-y-ăqɛm-Ø *yhwh* *šōp̄əṭ-īm* … *wə=**ḵī***
WAYQ-3M-establish-SG Yahweh judge-PL … and=**COMP**

hēqīm-Ø *yhwh* *lă-hɛm* *šōp̄əṭ-īm*
establish\PFV-3M.SG Yahweh for-them judge-PL

'And Yahweh established judges … *Whenever* Yahweh established judges for them, (Yahweh would be with the judge …)'

(4.44) 1 Sam. 14:27–29b: וַיָּשֶׁב יָדוֹ אֶל־פִּיו וַתָּאֹרְנָה[39] עֵינָיו: ... אֹרוּ עֵינַי כִּי טָעַמְתִּי מְעַט דְּבַשׁ הַזֶּה:

way-y-ɑ̊šɛḇ-Ø *yɑ̊ḏ-ō* *ʾɛl* *pī-w* *wat-t-ɑ̊ʾōr-ənɑ̊*
WAYQ-3M-return\CAUS-SG hand-his to mouth-his WAYQ-F-light_up-3PL
ʿēn-ɑ̊yw ... ʾōr-ū *ʿēn-ay* ***kī*** *ṭɑ̊ʿam-tī*
eye(F)-DU.his ... light_up\PFV-3PL eye-DU.mine **COMP** taste\PFV-1SG
məʿaṭ *dəḇaš* *haz=zɛ*
little.of honey the=this

'And [Jonathan] returned his hand to his mouth and his eyes lit up. (A soldier spoke up ..., "Your father swore ..., 'Cursed is the man who eats anything today'" (...) But Jonathan said: "My father has brought misfortune on the land. See that) my eyes lit up *when* I tasted a little of this honey!"'

In (4.42), it is clear from the context to Joseph (the Addressee) that his brothers (the Speaker) have been to their father. It is therefore assumed to be part of the Common Ground, and *kī* can be used to refer to this event. In (4.43) this is even clearer, as the fact that Yahweh established judges has been introduced just a few sentences before. Example (4.44) is similar: the event described by the *kī*-clause is introduced in v. 27, and v. 28 (only given in translation) shows that the Addressee (the soldier) is aware of it as well.

To conclude this subsection: there are two ways in which temporal and conditional *kī*-clauses interact with the Common Ground. As we just saw, in many cases the event or condition is already tracked by both the Speaker and the Addressee, independent of whether it is likely to occur or have occurred. More importantly, however, in all instances the event or condition is certain or likely to occur or have occurred. As a result, if the event or condition is not yet in the Common Ground, it is at least easily accommodated. In this way these cases support the hypothesis about the general discursive contribution of *kī*.

4.6.2 Resultative clauses

Resultative clauses (sometimes called "consequential clauses") describe the result of the event described in the matrix clause; in English, resultative *kī* can often be translated with (*so*) *that*. When the result is also the purpose for a volitional act described in the matrix clause, a resultative clause is quite similar to a causal clause:

39 Consonantal text: ותראנה *wat-t-irʾɛ-nɑ̊* WAYQ-F-see-3PL 'and [his eyes] saw'.

(4.45) Gen. 31:36: מַה חַטָּאתִ֣י כִּ֔י דָּלַ֖קְתָּ אַחֲרָֽי׃

mah ḥaṭṭāʾ-tī ***kī*** *ḏālaq-tā* *ʾaḥăr-āy*
what sin\PFV-1SG **COMP** chase\PFV-2M.SG behind-me

'How have I (Jacob) sinned *that* you (Laban) have chased after me?'

In (4.45), it is in the Common Ground that the Addressee chased after the Speaker, as the Addressee has just caught up with the Speaker when (4.45) is uttered. Note that there is no reason why a purpose or resultative clause in general should refer to a result in the Common Ground, as the following example with another resultative connective, לְמַעַן *ləmaʿan*, shows:

(4.46) Gen. 37:22: וַיֹּ֤אמֶר אֲלֵהֶם֙ רְאוּבֵ֔ן אַל־תִּשְׁפְּכוּ־דָ֖ם הַשְׁלִ֣יכוּ אֹת֑וֹ אֶל־הַבּ֤וֹר הַזֶּה֙ אֲשֶׁר־בַּמִּדְבָּ֔ר וְיָ֥ד אַל־תִּשְׁלְח֖וּ־בּ֑וֹ לְמַ֗עַן הַצִּ֤יל אֹתוֹ֙ מִיָּדָ֔ם לַהֲשִׁיבוֹ֖ אֶל־אָבִֽיו׃

way-y-ōʾmɛr-Ø *ʾălē-hɛm* *rəʾūḇēn* *ʾal* *t-išpəḵ-ū* *ḏām*
WAYQ-3M-say-SG to-them Reuben not 2-shed\JUSS-M.PL blood

hašlīḵ-ū *ʾōṯ-ō* *ʾɛl* *hab=bōr* *haz=zɛ* *ʾăšɛr* *b=am=miḏbār*
throw\IMP-M.PL OBJ-him to the=pit the=this REL in=the=desert

wə=yāḏ *ʾal* *t-išləḥ-ū* *ḇ-ō* ***ləmaʿan***
but=hand not 2-send\JUSS-M.PL against-him **in_order_to**

haṣṣīl *ʾōṯ-ō* *miy=yāḏ-ām* *la=hăšīḇ-ō* *ʾɛl*
rescue\INF OBJ-him from=hand-theirs to=return\CAUS.INF-him to

ʾāḇī-w
father-his

'But Reuben said to them, "Don't shed blood; throw him into this pit in the desert but do not stretch out your hand against them," *in order to* rescue him out of their hand to return him to his father.'

In (4.46), the fact that Reuben tries to save Joseph is not yet known to the Addressee (the reader). It can be accommodated based on the contents of the direct speech report, but even this is not always the case. For instance, in (4.47) with the purposive construction לְ *lə* 'to' + infinitive, there is no reason in particular to think that Laban should go shear his sheep at this moment:

(4.47) Gen. 31:19: וְלָבָ֗ן הָלַךְ֙ לִגְזֹ֣ז אֶת־צֹאנ֔וֹ

wə=lāḇān *hālaḵ-Ø* ***li=gzōz*** *ʾeṯ* *ṣōʾn-ō*
and=Laban go\PFV-3M.SG **to=shear\INF** OBJ flock-his

'And Laban had gone *in order to* (*lə* 'to' + infinitive) shear his sheep.'

Nevertheless, we do not find such cases with resultative *kī*. Even in cases with future reference time, the resultative clause refers to the Common Ground because the prospected result has already been discussed. For instance, in (4.48), though the resultative is an irrealis with future time reference, it refers directly back to Saul's proposal in the previous verse.

(4.48) 1 Sam. 18:18: מִי אָנֹכִי ... כִּי־אֶהְיֶה חָתָן לַמֶּלֶךְ׃

mī	*ʾănōḵī*	...	***kī***	*ʾ-ɛhyɛ*	*ḥăṯṯān*	*l=am=mɛlɛḵ*
who	I	...	**COMP**	1SG-be\IPFV	son_in_law	to=the=king

'(Saul said to David: "Here is my oldest daughter Merab; I want to give her to you in marriage ..." But David said to Saul:) "Who am I ... *that* I should be the king's son-in-law?"'

This supports the hypothesis that *kī* is marked for reference to the Common Ground.

4.6.3 Concessive clauses

With a concessive clause (English *though*) the Speaker concedes some information to the Addressee, but at the same time denies that this information is incompatible with the assertion made in the matrix clause:

(4.49) 2 Sam. 12:12: כִּי אַתָּה עָשִׂיתָ בַסָּתֶר וַאֲנִי אֶעֱשֶׂה אֶת־הַדָּבָר הַזֶּה נֶגֶד כָּל־יִשְׂרָאֵל

kī	*ʾattå*	*ʿåśī-ṯå*	*b=as=såṯɛr*	*wa=ʾănī*	*ʾ-ɛʿĕśɛ*	*ʾɛṯ*	
COMP	you	do\PFV-2M.SG	in=the=secret	and=I	1SG-do\IPFV	OBJ	
had=dåḇår	*haz=zɛ*	*nɛgɛḏ*	*kål*	*yiśrå'ēl*			
the=thing	the=this	before	all.of	Israel			

'*Though* you have acted in secret, I will do this before all of Israel.'

A concessive usually presupposes that the Addressee knows or can easily accommodate the conceded information. For instance, in English I can utter (4.50) only to someone of whom I know that they share my belief about France's likelihood to win; I have to assume that *France won't win this World Championship* is in the Common Ground or easily accommodated.

(4.50) *Though France won't win this World Championship, they are a treat to watch.*

The fact that concessive clauses always refer to Common Ground makes it pointless to compare *kī* with other markers of concessive clauses, such as אִם *ʾim* (Job 9:15; Jer. 15:1) or וְ *wə* (e.g. Jdg. 16:15). However, the fact that

the cases with *kī* all refer to information in the Common Ground or easily accommodated is of course entirely expected.

4.6.4 Low-frequency adverbial clauses: summary

This section reviewed four types of less frequent adverbial *kī*-clauses: temporal, conditional, resultative, and concessive clauses. It is particularly important that these lower frequency uses of *kī* adhere to the predicted patterns, since they are a priori less likely to have lexicalized and lost the [+distal] feature. The data reveals that in these *kī*-clauses the information provided in the clause is indeed always in the Common Ground, easily accommodated, or (rarely) imposed on the Common Ground for a discursive effect. This confirms the hypothesis that *kī* still has a [+distal] feature, which is interpreted as referring to the Addressee, and thus to the Common Ground.

4.7 Standalone *kī*-clauses

As is well-known and seen in the previous sections, *kī* usually connects two clauses. After classifying all instances of *kī* for the relation it establishes between the two clauses, some instances remained for which it is not clear that *kī* really connects two clauses. In the literature, these cases have often been referred to as "emphatic" or "asseverative", but in order not to make any assumptions I have classified them as "standalone". Reviewing these cases, it becomes clear that there are three ways in which standalone *kī* can be used: to introduce oaths (section 4.7.1), conducive and rhetorical polar questions (section 4.7.2), or exclamatives (section 4.7.3). It is true that each of these functions can reasonably be called "emphatic", but it is nevertheless valuable to make precise what kinds of emphasis can be provided by *kī*, exactly. By spelling out what types of "emphatic" interpretations there are, exactly, we can prevent this category from becoming a universal catch-all.

4.7.1 Oaths

Oaths can be described as sincere and earnest speech acts meant to assure the Addressee of a certain assertion or promise (Conklin 2011: 2). They are typically accompanied by what Conklin (2011: 13–30) calls an "authenticating element" that is meant to assure the Addressee of the Speaker's sincer-

ity. Oaths introduced by *kī* with two different authenticating elements are shown in (4.51–4.52):

(4.51) 2 Sam. 3:9ab: כֹּה־יַעֲשֶׂה אֱלֹהִים לְאַבְנֵר וְכֹה יֹסִיף לּוֹ כִּי כַּאֲשֶׁר נִשְׁבַּע יְהוָה לְדָוִד כְּי־כֵן אֶעֱשֶׂה־לּוֹ׃

kōh y-aʿăśɛ-Ø ʾĕlōhīm lə=ʾaḇnēr wə=ḵōh y-ōsīp̄-Ø
thus 3M-do\IPFV-SG God to=Abner and=thus 3M-continue\IPFV-SG

*l-ō **kī** ka=ʾăšer nišbaʿ-Ø yhwh lə=ḏāwiḏ **kī***
to-him **COMP** like=REL swear\PFV-3M.SG Yahweh to=David **COMP**

ḵēn ʾ-ɛʿĕśɛ l-ō
thus 1SG-do\IPFV for-him

'(And Abner said: "…) So do God to Abner, and even more — *that* as what Yahweh has promised David, *that* I do thus for him!"'

(4.52) 2 Sam. 12:5: חַי־יְהוָה כִּי בֶן־מָוֶת הָאִישׁ הָעֹשֶׂה זֹאת׃

*ḥay yhwh **kī** ḇɛn mămwɛṯ hă=ʾīš hă=ʿōśɛ-Ø zōʾṯ*
life.of[40] Yahweh **COMP** son.of death the=man the=do\PTCP-M.SG this

'By the life of Yahweh, *that* the man who does this is a dead man!'

There are 17 cases of standalone *kī* introducing an oath in my corpus, of which six refer to discourse-old information content.[41] For example, in (4.53), the previous verse has already made the question of when the people will stop pursuing their brothers a topic of discussion, and therefore tracked in the Common Ground:

(4.53) 2 Sam. 2:27a: וַיֹּאמֶר יוֹאָב חַי הָאֱלֹהִים כִּי לוּלֵא דִבַּרְתָּ כִּי אָז מֵהַבֹּקֶר נַעֲלָה הָעָם אִישׁ מֵאַחֲרֵי אָחִיו׃

*way-y-ōʾmɛr-Ø yōʾăḇ ḥay hă=ʾĕlōhīm **kī** lūlēʾ*
WAYQ-3M-say-SG Joab life.of[42] the=God **COMP** had_not

*dibbar-tå **kī** ʾăz mē=hab=bōqɛr*
speak\PFV-2M.SG **COMP** then from=the=morning

naʿălă-Ø hă=ʿām ʾīš mē=ʾaḥărē
go_up\MID.PFV-3M.SG the=people(M) one from=behind

ʾăḥī-w
brother-his

40 For the analysis of this form as a construct state, see Conklin (2011: 24–26).

41 1 Sam. 20:13; 25:34ab; 29:6a; 2 Sam. 2:27a; Ruth 1:17.

42 On this form see footnote 40.

'(Abner called out to Joab: "… How long won't you tell the people to return from after their brothers?") And Joab said: "By the life of God, ([I swear] *that*) had you not spoken, (*that*) then [only] from the morning onwards would the people have ceased [pursuing], each from behind his brother."'

In the remaining cases Common Ground cannot be accommodated either, but is imposed.[43] This is the case in (4.51–4.52) above. It is precisely the imposition that creates the interpretation as an oath. Forcing the Addressee to accept an assertion in the Common Ground, the Speaker effectively assures the Addressee of their own sincerity and commitment to this assertion, which is precisely what an oath does (Conklin 2011: 2).

Besides *kī*, a number of other particles can be used to introduce oaths. However, the most frequent of these (אִם *ʾim*) introduces conditional oaths (Conklin 2011: 31–45) and is not comparable to oaths with *kī*. There are a number of other ways to introduce oaths (Conklin 2011: 60–65), but they do not appear often enough to enable a comparison with *kī*. We can compare oaths with *kī* to other strong assertions, however. In the following fragment, the Speaker (Joseph) asserts up to three times that the Addressees (his brothers) are spies:

(4.54) Gen. 42:9–16: וַיֹּאמֶר אֲלֵהֶם מְרַגְּלִים אַתֶּם לִרְאוֹת אֶת־עֶרְוַת הָאָרֶץ בָּאתֶם׃

way-y-ōʾmɛr-Ø	*ʾălē-hɛm*	*m-əraggəl-īm*	*ʾattɛm*	*li=rʾōṯ*	*ʾɛṯ*
WAYQ-3M-say-SG	to-them	PTCP-spy-M.PL	you	to=see\INF	OBJ
ʿɛrwa-ṯ	*hå̄=ʾå̄rɛṣ*	*bå̄ʾ-ṯɛm*	…		
weakness-of	the=land	come\PFV-2M.PL	…		

'(And Joseph saw his brothers, … he said, "Where do you come from?" And they said: "From the land of Canaan, to buy grain for food.") …And he said to them: "Spies is what you are; to check out the weakness of the land you have come."

וַיֹּאמֶר אֲלֵהֶם לֹּא כִּי־עֶרְוַת הָאָרֶץ בָּאתֶם לִרְאוֹת׃

way-y-ōʾmɛr-Ø	*ʾălē-hɛm*	*lō'*	***kī***	*ʿɛrwa-ṯ*	*hå̄=ʾå̄rɛṣ*
WAYQ-3M-say-SG	to-them	no	**COMP**	weakness-of	the=land
bå̄-ʾṯɛm	*li=rʾōṯ*	…			
come\PFV-2M.PL	to=see\INF	…			

(But they said to him: "No, my lord! … We are all the sons of one man; we are honest men! …") But he said to them: "No, *for* you have come to check out the weakness of the land!"

43 1 Sam. 14:39b; 20:3b; 20:12; 26:10, 16; 2 Sam. 3:9ab, 35; 12:5; 15:21ab.

וַיֹּאמֶר אֲלֵהֶם יוֹסֵף הוּא אֲשֶׁר דִּבַּרְתִּי אֲלֵכֶם לֵאמֹר מְרַגְּלִים אַתֶּם׃

way-y-ōʾmɛr-Ø *ʾălē-hɛm* *yōsēp* *hūʾ* *ʾăšɛr* *dibbar-tī* *ʾălē-ḵɛm*
WAYQ-3M-say-SG to-them Joseph it REL speak\PFV-1SG to-you

lē=ʾmōr *m-əraggəl-īm* *ʾattɛm* ...
to=say\INF PTCP-spy-M.PL you ...

(They said: "... We are the sons of one man in the land of Canaan; the youngest is with our father at the moment, and one is no longer alive.") But Joseph said to them: "It is as I said, saying 'you are spies!'

זֹאת תִּבָּחֲנוּ חֵי פַּרְעֹה כִּי מְרַגְּלִים אַתֶּם׃

wə=y-ibbåḥăn-ū *diḇr-ē-ḵɛm* *ha=ʾĕmɛṯ* *ʾitt-əḵɛm*
and=3M-test\MID.IPFV-PL word(M)-PL-yours Q=true with-you

wə=ʾim *lō'* *ḥ-ē* *pārʿōh* ***kī*** *m-əraggəl-īm* *ʾattɛm*
and=if not life-of Pharaoh **COMP** PTCP-spy-M.PL you

(... send one of you to take your brother, while the rest of you is in prison.) And so your words will be tested, whether they are true — and if not, by the life of Pharaoh, ([I swear] *that*) you are spies!'"

Note that the first time that Joseph asserts that his brothers are spies, the clause is unmarked. The second time is marked by *kī*, but this *kī* is probably causal-adversative. The third time is a direct quotation of the first, and therefore does not contain *kī*, but the fourth is clearly an oath with *kī*, preceded by the authenticating element 'life of Pharaoh'.

We can explain the use of *kī* in two ways. On the one hand, the question of whether his brothers are spies is clearly under discussion in the Common Ground. We could therefore argue that *kī* marks this reference to Common Ground. On the other hand, we could also note that there is an increased level of certainty that Joseph wishes to express: if his brothers fail the test, how can they still claim to be honest men? This allows Joseph to impose the assertion on the Common Ground in the form of an oath. Of course, regular declarative sentences (like the first 'spies is what you are') can also be used when the information content is already in the Common Ground. Nevertheless, patterns like the one in (4.54) lend further support to the hypothesis that *kī* is marked for the use of Common Ground.

Conklin's explanation for the use of *kī* to mark oaths is that *kī* is a remnant of an originally longer formula (*I swear that ...*) after elision of the pred-

icate (2011: 59).[44] While this is possible, some questions remain. For example, it is unclear why the complementizer would not have been elided together with the predicate in oath formulas, with only intonation serving to distinguish the exclamative from a declarative (as in some rhetorical questions, like *I said something funny?*, and interrogatives more generally in languages like Italian). The theory proposed here provides an explanation why the complementizer was retained: it is crucial to oaths that they impose information on the Common Ground, and this aspect is marked by the [+distal] complementizer.

4.7.2 Conducive and rhetorical questions

There are five instances in my corpus of conducive and rhetorical questions using the polar interrogative marker הֲ *hă* followed by *kī*, two of which are negated (הֲלֹוא כִי *hă-lōʾ kī*).[45] In one instance (2 Sam. 9:1), *hă-k̲ī* introduces a regular question which happens to nominalize a clause of which the content is under discussion in the Common Ground;[46] this instance is excluded here and classified as a subject clause. I will rely on Moshavi (2009) for my analysis of the five instances under discussion here; references to theoretical linguistic work on these types of questions can be found in her article.

A rhetorical question has the form of an interrogative but is, at the discursive level, an implicit assertion rather than a request for information (Moshavi 2009: 32). The implicit assertion contributed by a rhetorical polar question is the negation of its propositional content: *Are you the president?* implies you are not the president (Moshavi 2009: 33). A conducive question is similar to a rhetorical question in that the Speaker has a certain prior belief regarding the correct answer, and may not expect an answer, but do not function as implicit assertions; for example, *Is that you, Henry?* does not im-

44 Note that a similar account has been proposed for English exclamatives with *that* (*I am surprised that he is still not here* > *That he is still not here!*; cf. Ross 1970). While there may be some diachronic truth to this account, it is unlikely that the elided predicate still plays a role synchronically. The reason for this is that if the underlying structure of an exclamative like *That he is still not here!* is *I am surprised that he is still not here*, there is no way to rule out that the actual underlying structure is *I tell you that I am surprised that he is still not here*, and so on ad infinitum (cf. Speas & Tenny 2003: 338 and chapter 3, footnote 13 for more discussion).

45 Without *lōʾ*: Gen. 27:36; 29:15; 2 Sam. 23:19. With *lōʾ*: 1 Sam. 10:1; 2 Sam. 13:28.

46 Similar to one more example outside my corpus, Job 6:22.

ply that you are Henry but merely conveys an expectation (Moshavi 2009: 38). Moshavi describes the discursive functions of conducive questions in Biblical Hebrew as (a) confirming a belief of the Speaker, (b) expressing surprise, (c) showing the Addressee that the Speaker knows something to be true, and (d) drawing attention to a fact (2009: 38 n. 38).

The questions with *hă-ḵī* 'is [it] that' are conducive, implying that the information content of the *kī*-clause is true:

(4.55) Gen. 27:36: הֲכִי֩ קָרָ֨א שְׁמ֜וֹ יַעֲקֹ֗ב וַֽיַּעְקְבֵ֙נִי֙ זֶ֣ה פַעֲמַ֔יִם

hă=ḵī *qårå̄ʾ-Ø* *šəm-ō* *yaʿăqōḇ* *way-y-aʿqəḇ-Ø-ēnī*
Q=COMP call\PFV-3M.SG name(M)-his Jacob WAYQ-3M-deceive-SG-me

zɛ *pāʿăm-ayim*
this time-DU

'*Isn't* his name Jacob?[47] He has deceived me these two times!'

In the case of conducive questions with *kī*, the information content is not only implied but also well-known to be true, and thus in the Common Ground. Compared to conducive questions without *kī*, the questions with *kī* convey a much stronger belief with respect to the expected answer. In (4.56) without *kī* the Speaker is much less certain that Saul should now be considered a prophet than that the Speaker in (4.55) is certain about Jacob's name. This explains why *kī* can be used in (4.55) but not (4.56): only in (4.55) can the propositional content be assumed to be in the Common Ground. In (4.56), the Speaker does not even want to impose it on the Common Ground.

(4.56) 1 Sam. 10:11: וַיְהִ֗י כָּל־יֽוֹדְעוֹ֙ מֵאִתְּמ֣וֹל שִׁלְשֹׁ֔ם וַיִּרְא֕וּ וְהִנֵּ֥ה עִם־נְבִאִ֖ים נִבָּ֑א
וַיֹּ֨אמֶר הָעָ֜ם אִ֣ישׁ אֶל־רֵעֵ֗הוּ מַה־זֶּה֙ הָיָ֣ה לְבֶן־קִ֔ישׁ הֲגַ֥ם שָׁא֖וּל בַּנְּבִיאִֽים׃

wa-y-əhī-Ø *kål* *yōḏəʿ-Ø-ō* *mē=ʾittəmōl šilšōm*
WAYQ-3M-be-SG all.of know\PTCP-M.SG-him from=before

way-y-irʾ-ū *wə=hinnē* *ʿim* *nəḇiʾ-īm* *nibbå̄ʾ-Ø*
WAYQ-3M-see-PL and=behold with prophet-PL prophesy\PTCP-M.SG

way-y-ōʾmɛr-Ø *hå̄=ʿå̄m* *ʾīš* *ʾɛl* *rēʿ-ēhū* *mah* *zɛ*
WAYQ-3M-say-SG the=people(M) one to companion-his what this

hå̄yå̄-Ø *lə=ḇɛn* *qīš* ***hă**=ḡam* *šå̄ʾūl* *b=an=nəḇīʾ-īm*
be\PFV-3M.SG to=son.of Kish **Q**=also Saul in=the=prophet-PL

47 The name Jacob translates as 'he-will-deceive'. Note that my translation uses a rhetorical question, and thus adds negation. Another option is "No wonder his name is Jacob" (NLT; cf. also NET). One can also take this *kī* as causal ('Is it *because* his name is Jacob that …'), which is also possible in Gen. 29:15 but not in 2 Sam. 23:19.

'And when all who knew him (Saul) from before saw how he prophesied with prophets, the people said to each other: "What happened to the son of Kish? *Is* Saul also among the prophets?"'

The questions with *hă-lōʾ kī* 'isn't [it] that' are rhetorical. Due to double negation (once for *lōʾ* 'not' and once for the rhetorical question), these questions also imply that the information content of the *kī*-clause is true:

(4.57) 1 Sam. 10:1: וַיִּקַּח שְׁמוּאֵל אֶת־פַּךְ הַשֶּׁמֶן וַיִּצֹק עַל־רֹאשׁוֹ וַיִּשָּׁקֵהוּ וַיֹּאמֶר הֲלוֹא כִּי־מְשָׁחֲךָ יְהוָה עַל־נַחֲלָתוֹ לְנָגִיד׃

way-y-iqqaḥ-Ø *šəmūʾēl* *ʾɛṯ* *paḵ* *haš=šɛmɛn* *way-y-iṣōq-Ø*
WAYQ-3M-take-SG Samuel OBJ flask.of the=oil WAYQ-3M-pour-SG

ʿal *rōʾš-ō* *way-y-išš̊q-Ø-ēhū* *way-y-ōʾmɛr-Ø* ***hă=lōʾ*** ***kī***
on head-his WAYQ-3M-kiss-SG-him WAYQ-3M-say-SG **Q=not** **COMP**

məšš̊ḥ-Ø-ăḵ̊ *yhwh* *ʿal* *naḥălṯ-ō* *lə=nš̊ḡīḏ*
anoint\PFV-3M.SG-you.OBJ Yahweh over inheritance-his for=leader

'Then Samuel took a flask of oil and poured it on his (Saul's) head, and he kissed him and said: "*Has* Yahweh *not* anointed you as leader over his inheritance?"'

In the other rhetorical question, Common Ground is imposed by making an implicit assertion (2 Sam. 13:28). A rhetorical question does not leave room for the Addressee to reject the implied assertion, which has the effect of imposing it on the Common Ground. Though the use of *kī* is not necessary to form these types of questions (see Moshavi 2009 for many examples without *kī*), it is still well-suited because of this interaction with the Common Ground.

4.7.3 Exclamatives

The third way in which standalone *kī* is used is to form exclamatives.[48] Exclamatives are sentences that "express the speaker's affective response to

48 This includes many examples traditionally analyzed as "emphatic" or "asseverative". I analyze these as exclamatives, which can be defined more precisely based on the theoretical linguistic literature. Though some scholars have been skeptical towards the existence of emphatic/asseverative *kī* (e.g. Bandstra 1982; Aejmelaeus 1986; Locatell 2017), most authorities still assume this notion is useful in at least some instances (e.g. Waltke & O'Connor 1990: §40.2.2b; Joüon & Muraoka 2006: §164; Miller 2003: 103–116; Holmstedt 2010: 85, 92). Schoors (1981), following Gordis (1943) and Dahood (1965, 1966, 1968, 1970) to some extent, sees many cases of emphatic *kī* in poetry. I will have nothing to say about this issue, since this study is limited to narrative texts.

a situation" (Michaelis 2001: 1039). As with conducive and rhetorical questions, a proposition can be recovered from an exclamative: *How nice weather it is!* implies that it is nice weather. Crucially, the Speaker of an exclamative assumes this propositional content to be in the Common Ground (Zanuttini & Portner 2003): *How nice weather it is!* cannot be used to convince the addressee that it is nice weather. This distinguishes exclamatives from both declarative sentences, which make no such assumption, and rhetorical questions, which impose information on the Common Ground rather than assuming it is already shared with the Addressee. Nevertheless, this reference to Common Ground makes *kī* a natural element to mark exclamatives (cf. section 3.3.1).

In the examples found in the Hebrew Bible, exclamatives with *kī* express shock/outrage (4.58),[49] remorse (4.59), or commitment (4.60) with respect to the recoverable proposition, though there is no reason to think other emotions could not be expressed this way as well.

(4.58) Gen. 18:20ab: זַעֲקַ֥ת סְדֹ֖ם וַעֲמֹרָ֑ה כִּי־רָבָ֕ה וְחַטָּאתָ֖ם כִּ֥י כָבְדָ֖ה מְאֹֽד׃

zaʿăqa-ṯ	*səḏōm*	*wa=ʿămōrå̄*	***kī***	*råbb-å̄*
outcry(F)-of	Sodom	and=Gomorrah	**COMP**	be_great\PFV-3F.SG

wə=ḥaṭṭå̄ʾṯ-å̄m	***kī***	*ḵå̄ḇəḏ-å̄*	*məʾōḏ*
and=sin(F)-theirs	**COMP**	be_heavy\PFV-3F.SG	very

'*That* the outcry of/concerning Sodom and Gomorrah is so great! And *that* their sin is so heavy!'

(4.59) Jdg. 10:10: וַיִּזְעֲקוּ֙ בְּנֵ֣י יִשְׂרָאֵ֔ל אֶל־יְהוָ֖ה לֵאמֹ֑ר חָטָ֣אנוּ לָ֔ךְ וְכִ֤י עָזַ֙בְנוּ֙ אֶת־אֱלֹהֵ֔ינוּ

way-y-izʿăq-ū	*bən-ē*	*yiśrå̄ʾēl*	*ʾɛl*	*yhwh*	*lē=ʾmōr*
WAYQ-3M-call_out-PL	son-PL.of	Israel	to	Yahweh	to=say\INF

ḥå̄ṭå̄ʾ-nū	*l-å̄ḵ*	*wə=**ḵī***	*ʿå̄zaḇ-nū*	*ʾɛṯ* *ʾĕlōhē-nū*
sin\PFV-1PL	to-you	and=**COMP**	abandon\PFV-1PL	OBJ God-ours

'And the Israelites called out to Yahweh, saying: "We have sinned! And *that* we have abandoned our God!"'

(4.60) 1 Sam. 10:24: וַיֹּ֤אמֶר שְׁמוּאֵל֙ אֶל־כָּל־הָעָ֔ם הַרְאִיתֶ֕ם אֲשֶׁ֥ר בָּחַ֖ר־בּ֣וֹ יְהוָ֑ה כִּ֥י אֵ֖ין כָּמֹ֥הוּ בְּכָל־הָעָ֖ם

49 Also 1 Sam. 17:28. We can also include Gen. 45:26a here, if we assume that the brothers are in part speaking to each other. This explanation is not ideal, but I have no better alternative at this point.

way-y-ōʾmɛr-Ø *šəmūʾēl* *ʾɛl* *kål* *hå̄=ʿå̄m* *har=rəʾī-ṯɛm* *ʾăšɛr*
WAYQ-3M-say-SG Samuel to all.of the=people Q=see\PFV-2M.PL REL

bå̄ḥar-Ø *b-ō* *yhwh* ***kī*** *ʾēn* *kå̄mō-hū*
choose\PFV-3M.SG in-him Yahweh **COMP** not_exists like-him

bə=ḵål *hå̄=ʿå̄m*
in=all.of the=people

'Then Samuel said to all the people: "Have you seen whom Yahweh has chosen? *That* there is no one like him (David) among all the people!"'

These examples refer to Common Ground in the following way. First, in (4.58), God speaks to Abraham. Abraham has already had previous contact with Sodom in Genesis 13–14, where he has been able to see that the Sodomites are wicked sinners (13:13). This allows the Speaker to presuppose this information, and thus permits the interpretation as an exclamative. The use of Common Ground is even clearer in (4.59), where the Addressee (Yahweh) must know that the Israelites have abandoned him. Finally, I classified (4.60) as accommodated information content: the people can infer, based on the information that Yahweh has chosen David, that there is no one better suited for the task.

It is also possible to express commitment towards a wish, which has the effect of strengthening a wish (4.61).[50] It should be noted that in these cases the fact that the Speaker has a certain desire should already be in the Common Ground: in (4.61), it is already clear to the Addressee that the Speaker wants the spoil to be divided equally. It thus does not appear to be possible to strengthen just any wish with *kī*.

(4.61) 1 Sam. 30:24: כִּ֣י כְחֵלֶק֩ הַיֹּרֵ֨ד בַּמִּלְחָמָ֜ה וּכְחֶלֶק֙ הַיֹּשֵׁ֣ב עַל־הַכֵּלִ֔ים יַחְדָּ֖ו יֶחֱלֹ֑קוּ׃

kī *kə=ḥēlɛq* *hay=yōrēḏ-Ø* *b=am=milḥå̄må̄*
COMP like=part.of the=go_down\PTCP-M.SG in=the=battle

ū=ḵə=ḥēlɛq *hay=yōšēḇ-Ø* *ʿal* *hak=kēl-īm* *yaḥdå̄w*
and=like=part.of the=sit\PTCP-M.SG on the=item-PL together

y-aḥălōq-ū
3M-divide\IPFV-PL

'("Since they didn't go with me, we will not give them from the spoil. ..." But David said: "No! ...) *That* as the part of he who goes down in battle, so be the part of he who remains with the equipment! Together they shall divide it."'

50 Also 1 Sam. 8:9; 14:44 (if not an oath); 25:28a.

Exclamatives are clearly emphatic, but it is a much narrower category than "emphatic" or "asseverative" *kī*. Before classifying an instance of *kī* as an exclamative it must be shown that the propositional content is in the Common Ground or easily accommodated by the Addressee.[51]

4.8 Conclusion

In this chapter I have argued that the Biblical Hebrew particle *kī* has inherited a [+distal] feature, which can be interpreted on the discourse level in various ways. Contrary to previous accounts, I propose that this [+distal] feature is not only relevant for a historical description of the development of the various functions of this particle, but is still present around the time of Biblical Hebrew. This greatly simplifies the description of the semantic-pragmatic contribution of this particle.

Though distancing a clause can have the effect of backgrounding its information content, most commonly it marks the information content as belonging to the Common Ground.[52] The reason for this is that the Common Ground involves the Addressee, who is seen as distal from the Speaker. A comparison of *kī* with other grammatical strategies performing similar functions shows that *kī* is used in particular when the information content in the *kī*-clause is already in the Common Ground or can easily be accommodated by the Addressee. In some cases, this function is used by the Speaker to im-

51 I therefore do not support Schoors' (1981: 251) claim that adversative *kī* is a special kind of asseverative clause: contrary to standalone *kī*, adversative *kī* links two clauses, and the latter also appears far more often without reference to Common Ground (see section 4.5). I also find it unlikely that exclamative *kī* developed from the use in oaths, as suggested by (Williams 2007: §449), since oaths impose content on the Common Ground, which is a less basic operation than assuming that content is in the Common Ground as exclamative *kī* does. The reverse development seems to be more likely.

52 Follingstad (2001) attempts to reduce *kī* to a "discourse deictic particle" with the function of setting up a new Mental Space and shifting viewpoint to this space. This can perhaps be seen as a form of "distancing", and in some instances, Follingstad recognizes reference to Common Ground (e.g., *kī* is said to be able to "'point to' a previously explicitly stated proposition"; 2001: 152). However, it is not clear that the Mental Space Theory is precise enough to exclude uses of *kī* that do not occur, and in some cases it actually makes incorrect predictions. For example, Follingstad (2001: 268–269) claims that *kī* can introduce a "hypothetical" conditional to which the Speaker does not need to commit, while the Speaker does need to commit to the truth of conditionals introduced by אִם *'im* 'if'. This is contrary to the consensus, which I have supported in section 4.6.2.

pose information on the Common Ground, with various discursive effects (e.g., assuring the Addressee of a certain assertion in an oath). The various functions that scholars have identified as being expressed by *kī* can, with only two exceptions, be seen as contextually derived from this general notion of Common Ground.

Exceptions to this general pattern mostly occur in the most common function, causal *kī*: here we find instances that do not refer to Common Ground and cannot even be seen as backgrounding (section 4.4). I therefore conclude that some lexicalization must have occurred here, which subsequently transferred to the categories of causal-adversative and adversative *kī* (section 4.5). In all other functions, the vast majority of instances refer to the Common Ground. Though grammaticalization is required to account for the use of *kī* to introduce subject and object clauses (section 4.3), we do not need to assume lexicalization to account for the various adverbial uses of *kī* (section 4.6; contra e.g. Locatell 2017, 2020). These are more economically described as pragmatically inferred uses based on the general function of marking Common Ground; since some of these functions are very infrequent, assuming semantic shifts for which we do not have evidence is problematic. Taking all of this into account, I propose a description of *kī* with only three distinct functions:[53]

1. Referring to Common Ground (including easily accommodated and imposed information content)
    (a) As a complementizer introducing subject and object clauses
    (b) When connecting two clauses: introducing adverbials (adversative, causal, causal-adversative, concessive, conditional, resultative, temporal)[54]

53 Park (2016), like me, aims to reduce the description of *kī*, but uses the notion of "nominalization". In South Asian languages, nominalization constructions can have a wide variety of functions (Yap et al. 2011) that indeed show a curious overlap with the functions of *kī*. Unfortunately, the notion of "nominalization" is not very well-defined, and we must be careful not to simply replace one ill-defined notion ("emphasis", "asseveration") with another one. At the very least, an explanation should be given why "nominalizers" across unrelated languages take on similar functions, and what the relation between the form and function of these "nominalizers" is.

54 Given that some of these types are very infrequent, it is conceivable that there are still other types of adverbials that could be marked by *kī*, which have not made their way into the corpus. This is fine, as long as the information content is in the Common Ground and the discursive function can be inferred from context.

   (c) When standalone: introducing oaths, conducive and rhetorical questions (with הֲ *hă*), and exclamatives

2. Lexicalized causal meaning ('because', 'for', etc.)
3. Lexicalized adversative meaning ('but'), developed from causal via causal-adversative ('not X, *but*/*because* Y')

Many authors, Aejmelaeus (1986: 193–194) most verbosely, have wondered how native speakers could have distinguished between the many uses of *kī*. My answer to this question is as follows. The most common and default function of *kī* is to mark Common Ground. When the Addressee cannot accommodate the information content of a *kī*-clause, this triggers them to use one of the lexicalized meanings instead; choosing between causal and adversative is possible based on context. In its default function of marking Common Ground, one of the three subfunctions can be selected based on simple cues: the existence of a matrix predicate or the place of the *kī*-clause in subject position for complementizer *kī*, authenticating elements for oaths, the interrogative particle for questions, and intonation for exclamatives; anything else is adverbial. The appropriate kind of adverbial meaning can be selected based on context.

Part III

Authority and dignity

Liṗnē 'in the face of'

A Locative preposition with a threatening connotation

5

Abstract The Biblical Hebrew expression *liṗnē*, literally 'to the face of', has grammaticalized into a Locative preposition 'before'. Its function in combination with the root *ngp* 'inflict, defeat' is not clear: what is the role of the Philistines in 'Israel was defeated *before* the Philistines' (1 Sam. 4:2)? Some scholars have used grammaticalization theory to argue that *liṗnē* is an Agent marker here: 'Israel was defeated *by* the Philistines'. However, this view is untenable in the face of arguments based on narrative structure, syntactic-semantic restrictions, grammaticalization theory, and language typology. I show that *liṗnē* is a simple Locative prepositional expression, but that the element 'face' has the connotation that Israel is threatened by the Philistines: Israel is in the 'realm of influence' of the Philistines. In present-day English, the near-literal translation 'in the face of' is a better alternative. Based on parallels in the active voice, I argue that Yahweh is the actual, implicit, Agent of *ngp* 'inflict, defeat'. In fact, English *before*, like *liṗnē*, used to have a connotation of threatening influence. In recent times, this meaning shifted so that the original translation became misunderstood.

5.1 Introduction

The expression *liṗnē*, formed by *lə* 'to' with the construct state of *pā̆nīm* 'face', usually functions as a preposition with a spatial or temporal meaning ('in

This chapter was originally published as: Camil Staps. 2022. *Liṗnē* 'in the face of': A Locative preposition with a threatening connotation. *Journal of Hebrew Scriptures* 22(1). 1–24. DOI: 10.5508/jhs29582.

I am grateful to Johan Rooryck, Ellen van Wolde, and the anonymous reviewers of the *Journal of Hebrew Scriptures* for their thoughtful comments on earlier versions of this article.

front of'; 'before').[1] It has a figurative sense when it marks an argument of the verb *ngp* 'inflict, defeat', and is then commonly translated either as 'before' ([N]KJV, A/R/ESV, NASB95) or as the Agent marker 'by' (NLT, NIV, NASB20).[2] Some examples follow:[3]

(5.1) 1 Sam. 4:2: וַיִּנָּגֶף יִשְׂרָאֵל לִפְנֵי פְלִשְׁתִּים

way-y-innɑ̊̄gɛp̄-Ø *yiśrɑ̊̄ʾēl* ***li=p̄n-ē*** *p̄əlištī-m*
WAYQ-3M-defeat\MID-SG Israel **to=face-of** Philistine-PL

'Israel was smitten *before* the Philistines.' (KJV)
'Israel was defeated *before* the Philistines.' (ESV)
'Israel was defeated *by* the Philistines.' (NIV)

(5.2) 2 Sam. 2:17: וַיְנֻגָּף אַבְנֵר וְאַנְשֵׁי יִשְׂרָאֵל לִפְנֵי עַבְדֵי דָוִד׃

way-y-innɑ̊̄gɛp̄-Ø *ʾaḇnēr* *wə=ʾanš-ē* *yiśrɑ̊̄ʾēl* ***li=p̄n-ē***
WAYQ-3M-defeat\MID-SG Abner and=man-PL.of Israel **to=face-of**

ʿaḇḏ-ē *ḏɑ̊̄wiḏ*
servant-PL.of David

'..., and Abner was beaten, and the men of Israel, *before* the servants of David.' (KJV)
'And Abner and the men of Israel were beaten *before* David's servants.' (ESV)
'..., and Abner and the Israelites were defeated *by* David's men.' (NIV)

For modern readers, the translation 'before' is hard to understand in this context, because in present-day English *before* has a purely spatial or temporal meaning (*the bill is presently before Congress*; *the day before yesterday*). Recent, freer translations have gone a step further in their interpretation

1 This type of expression is commonly called a "complex preposition". Joüon & Muraoka (2006: §1030) describe them in Biblical Hebrew as "pseudo-prepositions", by which they mean "a combination of one of the prepositions ... and a substantive, often lexemes denoting parts of the body ... in the status constructus". According to Waltke & O'Connor (1990: §11.3a), their meaning is often not predictable from the constituent components and these expressions "function syntactically as prepositions" (Waltke & O'Connor 1990: §11.3.1a). In the theoretical linguistic literature there is some debate as to how these expressions should be analyzed (Seppänen et al. 1994; Huddleston & Pullum 2002: 620–623; Hoffmann 2005; Pullum 2006). These issues have repercussions for the analysis of the Hebrew expressions as well, but this must be left for another occasion. For this chapter we can assume that these expressions are functionally, if not syntactically, prepositions. However, to avoid any confusion I will refer to them as "expressions", not prepositions.

2 I will use the capitalized terms Agent, Patient, Intermediary Agent, Instrument, Cause, Locative, and Source for thematic roles in the tradition of Gruber (1965). There is no

and use the Agent marker 'by', an analysis that finds support from Sollamo (2003), Rodriguez (2017: 180), and Jones (2018). Although translation to English is not the focus of this chapter, different choices reveal different interpretations of the underlying Hebrew construction, and thus show that this construction requires explanation. This will be the focus of this chapter.

In contrast to the recent tendency to interpret *lipnē* as an Agent marker, I will show here that the meaning of *lipnē* and *before* in cases like (5.1–5.2) is actually Locative. To explain its meaning in this context, I draw attention to a pragmatic connotation of the body part 'face' which expresses that the complement (e.g. the Philistines in [5.1]) has some threatening influence over the subject (Israel). This connotation can be approximated with 'in the face of' or 'in confrontation with' in present-day English. I show that *before* also had this connotation in older stages of English, which suggests that the traditional translations in (5.1–5.2) were perfectly intelligible at the time of writing. The problem is therefore not that the meaning of the He-

universally agreed upon list of thematic roles and their definitions, but the ones I use here are all relatively standard in theoretical linguistics. It is important, however, that they are understood as prototypes in the sense of Dowty (1991: 571–575), so that an argument can fill an Agent slot even if it is not in all aspects like the Proto-Agent. This is somewhat similar to the Actor and Undergoer macroroles in Role and Reference Grammar (Foley & Van Valin 1984), and the argument advanced in this chapter can be reformulated in that framework. For Dowty, the Proto-Agent is a *sentient* argument that *volitionally* brings about an *event* or *change of state in another participant* and *moves relative to the position of another participant*; finally it *exists independently of the event* (Dowty 1991: 572). The Proto-Patient *undergoes a change of state* and is *causally affected by another participant*; it is *stationary relative to movement of another participant* and *does not exist independently of the event or not at all* (Dowty 1991: 572 — Dowty's additional property *incremental theme* is not relevant to us here). The traditional roles can be defined using these properties (Dowty 1991: 577): the Agent has volition, causation, and sentience (and movement, but this is not relevant for our purpose). The Patient has change-of-state and causally-affectedness (stationarity and dependent existence are again not relevant here). An Instrument has causation without volition or sentience. I define the Intermediary Agent as an Agent with low volitionality. For Cause I rely on Palancar (2002), who defines this role as "the role played by a given entity — normally conceived of as either an abstract or natural force — construed as the causal force which has brought about a certain state of affairs" (Palancar 2002: 27). A Locative picks out a position in a region, which can be spatial (*The ball is in the box*) or temporal (*He was elected in 2008*); a Source refers to the point of origin of an event of motion. For modern discussion on thematic roles, see the reference works (e.g. Davis 2011; Harley 2011; Primus 2016).

3 An exhaustive list will be provided in table 5.1 on page 182.

brew expression *lip̄nē* is obscure and requires explanation (the path taken by Sollamo 2003, Rodriguez 2017: 180, and Jones 2018), but that older translations have become misunderstood due to changes in English. After elucidating these, there is no need for an Agent marker interpretation anymore. Below I also show why this interpretation is in itself unlikely regardless of any alternatives.

Before turning to the evidence for a connotation of threatening influence for Biblical Hebrew *lip̄nē*, let me clarify what I mean with this connotation by examining English *in the face of* and *in confrontation with*. English has phrases like *in the face of adversity/death/…*, where the complement has an unpleasant connotation by itself. By contrast, combinations with a complement with a beneficial or pleasant connotation are infelicitous or receive an ad hoc negative connotation: *#in the face of profit/the weekend/…*. Likewise, phrases that are of themselves neutral receive a negative connotation when used as the complement of these expressions: *in the face of change/…*; *in confrontation with the system/other communities/…*. Therefore, the prepositional expression provides a pejorative connotation if its complement does not have one already. The complement has a threatening influence in the sense that it has the ability to negatively influence another entity.

What might cause this connotation? It is striking that *in the face of* and *in confrontation with* both incorporate a word for 'face'.[4] In many languages, words for 'face' are related to concepts of authority and dignity. We find evidence for this in English (*lose/save/retain face*) and many languages around the world, as the following examples demonstrate:[5]

(5.3) Jordanian Arabic (Al-Adaileh & Abbadi 2012: 81)
akal widʒhi 'he has eaten my face': 'he has harassed me'

(5.4) Turkish (Kraska-Szlenk 2014: 30)
yüz akı 'clear face': 'honor'

(5.5) Swahili (Kraska-Szlenk 2014: 30)
sina uso 'I have no face': 'I am ashamed'

4 Observe that *confrontation* comes from Latin *frōns* 'forehead'.

5 However, one must be careful not to generalize here (see e.g. Littlemore 2019: 192–201 on cultural variation in embodied metaphor; cf. Strecker 1993 on the body part face in particular). All these languages show that the body part face is in *some* way related to concepts of authority and dignity, but the exact range in which these metaphors can be used will vary.

(5.6) Chinese (Yu 2001: 16)

a. *diu lian* 'lose face': 'lose face, be disgraced'

b. *mei lian* 'no face': 'feel ashamed, feel embarrassed'

c. *yao lian* 'want face': 'care about one's reputation'

(5.7) Thai (Ukosakul 2005: 119–120)

a. *ráksă nâ* 'preserve face': 'preserve someone's ego'

b. *mâi hâi*/*wái nâ* 'not giving/keeping face': 'too direct, inconsiderate'

Kraska-Szlenk (2014: 30) recognizes the same relationship in human behaviour, in the fact that we hold our face up when we are proud and want to seem respectable while we bow our heads low when we are ashamed. This observation finds support in the Hebrew Bible itself, when Cain's face 'falls' (*npl*) when his offer is rejected by Yahweh (Gen. 4:5–6).[6] The combination of cross-linguistic, psychological, and biblical evidence shows that these words are not arbitrarily related, but that this phenomenon reflects a psychological reality.

The 'face' is not just a metaphor for concepts like authority and dignity, it also comes to denote the *space* in which a person can exercise authority or enforce their dignity.[7] This is clear from the many idioms where a word for 'face' is combined with a spatial preposition (5.8). These expressions stem from the unwanted intrusion upon someone's sphere of authority and dignity.

(5.8) a. *The voters are saying, "In your face, Bush!"*
(*N.Y. Times* 6 Jun. 1992, 23/1 in citation by OED: s.v. face, n., P5d(c))

b. *Fuck off, scumbag. Get out of my face.*
(Wilson 2005, *Cusp*, 51 in citation by OED: s.v. face, n., P5g(a))

c. Dutch (Den Boon & Hendrickx 2015–2017: s.v. gezicht, 7)
iemand in zijn gezicht uitlachen 'laugh in someone's face'

6 Compare also Job 29:24, where Job speaks of his former glory: '(the men around me) would not let the light of my face fall (hiphil *npl*)', i.e., they would pay Job respect in order to preserve his face.

7 Such an extension is common in spatial expressions. For example, *front* normally denotes a vertically oriented bounded region on an object (e.g. the *front* of a house), but this region can be projected onto the surrounding (horizontally oriented) environment with *in front of* (Jackendoff 1996: 15). Similarly, words for 'face' can denote someone's authority and dignity, but also the projected region in which someone can exercise their authority or enforce their dignity.

All in all, the cognitive underpinning for a connotation of threatening influence for the body part 'face' seems clear: the face is related to concepts of authority and dignity, and comes to denote the abstract region of a person's authority and dignity. When you enter this realm of authority, two things can happen: either you challenge their authority, as in (5.8), or you relinquish some of your own authority and permit that person's influence over you. This is seen as a threat, which gives rise to the connotation of threatening influence for *in the face of* and *in confrontation with*.

Biblical Hebrew *pånīm* 'face' has very similar connotations.[8] Example (5.9) is highly reminiscent of the provocative *in your face* in (5.8a).[9] But being in someone's realm of authority does not necessarily imply friction, as (5.10) demonstrates: here, 'to stand to someone's face' has come to mean 'to serve them'.[10]

(5.9) Job 1:11: עַל־פָּנֶיךָ יְבָרְכֶךָּ׃

ʿal **pånɛ**-*ḵå* *y-əḇårăḵ-Ø-ɛkkå*

on **face**-yours 3M-bless\IPFV-SG-you.OBJ

'He will curse[11] you *in* your *face*.'

---

8 Rodriguez (2017: 179) already recognized a "metaphor of *face-personal space* as *dominance*" (emphasis original). The evidence I adduce for the connotation of threatening influence is slightly more general, as it also covers his "service metaphor", for example (Rodriguez 2017: 178). It is not exactly clear what Rodriguez' method is to distinguish these categories. We can also relate the metaphors with *pånīm* 'face' outlined below to the expression *ḥēn bəʿēnē* 'favour in the eyes of', which Vardi (2015) described as expressing a hierarchical relationship.

9 Also Job 2:5.

10 Similarly: Gen. 41:45; 1 Kgs. 18:15; Deut. 1:38; 1 Sam. 16:21, 22; 1 Kgs. 1:2; 10:8 (with *ʿaḇdɛḵå* 'your servants'); 12:6, 8; 17:1; 18:15; 2 Kgs. 3:14; 5:16; Jer. 15:19; 18:20; 52:12; Ezek. 8:11; Est. 4:5; Dan. 1:19. With *šrt* 'serve' besides *ʿmd* 'stand' we find Num. 16:9; Ezek. 44:11, 15. See also Ringgren's (2001: 182–183) categories "serve", "priestly service", and "worship" in his analysis of *ʿmd*. Even when *lip̄nē* seems purely Locative and *ʿmd* 'stand' is to be taken literally, the prepositional object is almost always the more dominant party: Gen. 18:22; 19:27; 43:15; 47:7; Exod. 9:10, 11; Lev. 27:8; Num. 3:6; 5:16, 18, 30; 27:2, 19, 21, 22; 35:12; Deut. 19:17; 29:14; Josh. 10:8; 20:6, 9; 21:44; 32:9; Jdg. 2:14; 1 Sam. 6:20; 1 Kgs. 1:28; 19:11; 2 Kgs. 4:12; 5:15; 8:9; Jer. 7:10; 15:1; 35:19; 40:10; 49:19; 50:44; Zech. 3:1, 3, 4; Est. 8:4; Dan. 2:2; Ezr. 9:15; 2 Chr. 9:7; 10:6, 8; 18:20. Two seemingly contradictory cases can be explained away: Exod. 17:6 (God before Moses, but God is helping Moses); Lev. 18:23 (a woman before an animal, but passively allowing it to have sex with her). We also find *pånīm* with *šrt* 'serve' without *ʿmd* 'stand': 1 Sam. 2:11, 18; 3:1; Est. 1:10; 1 Chr. 6:17; 16:4, 37.

11 The verb *brk* 'bless' can be used as a euphemism for 'curse'.

(5.10) Dan. 1:5: וּמִקְצָתָם יַעַמְדוּ לִפְנֵי הַמֶּלֶךְ׃

ū=mi=qṣå̱ṯ-å̱m *y-aʿamḏ-ū* ***li=p̄n-ē*** *ham=mɛlɛḵ*
and=from=end-theirs 3M-stand\IPFV-PL **to=face-of** the=king

'And at their end (of the years in training) they (the men in training) would serve (lit.: stand *to the face of*) the king.'

The threatening connotation of this authority is clearly visible in (5.11–5.13). In (5.11), *mippənē* 'from the face of' is not a neutral Source preposition, but expresses that Sarai has a kind of threatening influence over Hagar, her maid at whom she is angry. With this verb, a truly neutral Source is marked with *min* 'from'.[12] Mercy (*raḥămin*) is something you need 'in the face of' somebody who may otherwise do something to you (5.12).[13] The use of *mippənē* in the context of fear was already recognized by Rodriguez (2017: 194–195), but he seems to limit it to cases where fear is made explicit with a verb like *yrʾ* 'fear' (5.13).[14]

(5.11) Gen. 16:6: וַתְּעַנֶּהָ שָׂרַי וַתִּבְרַח מִפָּנֶיהָ׃

wat-t-əʿannɛ-Ø-hå̄ *śå̄ray* *wat-t-iḇraḥ-Ø* ***mip=på̄nɛ-hå̄***
WAYQ-F-oppress-3SG-her Sarai WAYQ-F-flee-3SG **from=face**-hers

'So Sarai oppressed her (Hagar), and she (Hagar) fled *from* her (Sarai's) *presence* (lit.: *face*).'

(5.12) Neh. 1:11: אָנָּא אֲדֹנָי ... וּתְנֵהוּ לְרַחֲמִים לִפְנֵי הָאִישׁ הַזֶּה

ʾå̄nnå̄ʾ *ʾăḏōn-å̄y* ... *ū=ṯən-Ø-ēhū* *lə=raḥămīm* ***li=p̄n-ē***
please lord-mine ... and=give\IMP-M.SG-him to=mercy **to=face-of**

hå̄=ʾīš *haz=zɛ*
the=man the=this

'My Lord, ... grant [your servant, i.e. the Speaker] mercy *before* this man.'

12 This can be seen in Isa. 48:20, where it marks the Chaldeans when they are no longer a threat to Israel. Apart from this example, *min* usually marks locations ('the land' 2 Sam. 19:9; 'afar' Isa. 22:3). But with *mippənē* the argument is always threatening: Gen. 35:1, 7; Exod. 2:15; Jdg. 11:3; 1 Sam. 21:11; 1 Kgs. 2:7; 12:2; Ps. 3:1; 57:1; 2 Chr. 10:2. The nuance of *millipnē* as opposed to *mippənē* is unclear: Jon. 1:3. We also find a few other complex prepositions with *min*: *mēʿal* 'from upon' (2 Sam. 19:9 — paralleled by *mippənē* in 1 Kgs. 2:7; Neh. 13:28); *mēʾēṯ* 'from with' (1 Kgs. 11:23). These do not occur frequently enough to be able to discuss them here.

13 Similarly: Gen. 43:14; 1 Kgs. 8:50; Ps. 106:46; Dan. 1:9; 2 Chr. 30:9.

14 Also Exod. 9:30; Deut. 5:5; Josh. 9:24; 11:6; 1 Sam. 7:7; 18:29; 21:13; 1 Kgs. 3:28; 2 Kgs. 1:15; 19:6; 25:26; Isa. 37:6; Jer. 1:8; 41:18; 42:11; Hag. 1:12; Neh. 4:8. With *millipnē*: 1 Sam. 18:12; Eccl. 8:12.

(5.13) 1 Kgs. 1:50: וַאֲדֹנִיָּ֣הוּ יָרֵ֔א מִפְּנֵ֖י שְׁלֹמֹ֑ה

wa=ʾăḏōniyyåhū yårēʾ-Ø ***mip=pən-ē*** *šəlōmōh*
and=Adonijah fear\PFV-3M.SG **from=face-of** Solomon

'And Adonijah feared (*because of*) Solomon.'

The use of *ʿal pənē* 'on the face of' in (5.14) does not mean simply 'alongside' but underlines the enmity between Ishmael and his kinsmen:[15]

(5.14) Gen. 16:12: וְהוּא֙ יִהְיֶ֣ה פֶּ֔רֶא אָדָ֖ם יָד֣וֹ בַכֹּ֑ל וְיַ֤ד כֻּל־בּוֹ֙ וְעַל־פְּנֵ֣י כָל־אֶחָ֔יו יִשְׁכֹּֽן׃

wə=hūʾ y-ihyɛ-Ø pɛrɛʾ ʾåḏåm yåḏ-ō ḇ=ak=kōl
and=he 3M-be\IPFV-SG wild_ass.of man hand-his against=the=all

wə=yaḏ kōl b-ō ***wə=ʿal pən-ē*** *ḵål ʾɛḥ-åyw*
and=hand.of all against-him **and=on face-of** all.of brother-PL.his

y-iškōn-Ø
3M-live\IPFV-SG

'And he (Ishmael) will be a wild ass of a man; his hand will be against everyone and everyone's hand will be against him; and *in the face of* all his brothers he will live.'

Finally, (5.15) shows that a threatening connotation can also be present when *pånīm* is used without any preposition:[16]

(5.15) Gen. 43:3: הָעֵ֤ד הֵעִד֙ בָּ֣נוּ הָאִ֔ישׁ לֵאמֹ֖ר לֹֽא־תִרְא֣וּ פָנַ֔י בִּלְתִּ֖י אֲחִיכֶ֥ם אִתְּכֶֽם׃

hå̄ʿēḏ hēʿiḏ-Ø ḇå̄-nū hå̄=ʾīš lē=ʾmōr lō'
warn\INFABS warn\PFV-3M.SG to-us the=man to=say\INF not

ṯ-irʾ-ū ***på̄n-ay*** *biltī ʾăḥī-ḵɛm ʾitt-əḵɛm*
2-see\IPFV-M.PL **face**-mine unless brother-yours with-you

'The man (Joseph) has sternly warned us: "You will not see my *face* (i.e., enjoy my powerful presence) unless your brother is with you."'

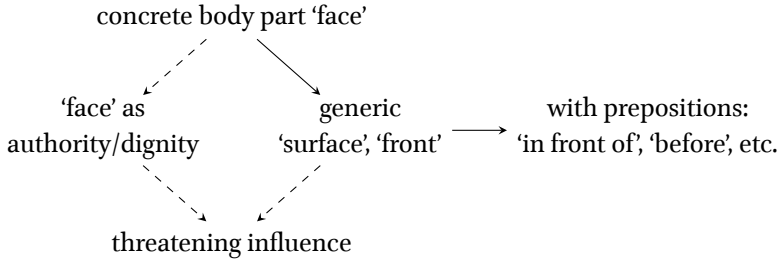
The extensions of Biblical Hebrew *pånīm* relevant here are schematized in the partial semantic-pragmatic network in (5.16). In this diagram, solid arrows represent semantic extensions (through grammaticalization), and dashed arrows represent pragmatic connotations. On the left we see the relationship with concepts of authority and dignity, as seen in (5.9–5.10). At

15 As rendered by some newer translations: 'in hostility toward all his brothers' (NIV); 'in defiance of all his brothers' (NASB20). Since *ʿal pənē* can also mean 'to the east of' (e.g. 1 Sam. 15:7; Kimron 1980), this verse probably contains a short origin myth explaining why the Ishmaelites live to the east of Israel. See also Gen. 25:18.

16 This example is due to Rodriguez (2017: 179), who also cites Gen. 43:5 and Exod. 10:28.

the same time, the body part undergoes semantic bleaching and becomes a generic noun for 'surface' or 'front'. A similar process can be seen in words like *facade* and *surface*, as well as in the use of *face* for 'surface' in *darkness was over the face of the deep* (Gen. 1:2 ESV). The combination of these two developments allows the interpretation of *pānīm* as the 'realm' of authority and dignity, which, as explained above, leads to the connotation of threatening influence seen in (5.11–5.15). The combination of *pānīm* with prepositions (*lipnē* 'to the face of'; *mippənē* 'from the face of'; etc.), illustrated on the far right, is quite independent from this connotation and only depends on the generic meaning 'surface', 'front'. For instance, *lipnē* describes a location in a region described by *pānīm*; *mippənē* describes the movement away from a location in that region; etc. By virtue of compositionality the connotation is still present in these derived expressions, but the connotation does not depend on the combination with a preposition.

(5.16) Semantic and pragmatic extensions of Hebrew *pānīm* 'face'.



I purposefully combine the noun *pānīm* and prepositions built on this noun in one diagram in (5.16). The reason for this is that metaphorical extensions like that of threatening influence are found with more than one prepositions (5.9–5.14) as well as without any preposition (5.15). Therefore, the extension is part of the meaning of *pānīm* and did not occur as a result of the grammaticalization of these prepositional expressions. The fact that the extension is also present in these derived expressions indicates that it must have already been part of the meaning of *pānīm* at the moment that these derived expressions became lexicalized. The alternative explanation, that the connotation developed independently in each grammaticalized preposition, is less economical.[17] This is not to say that the expressions did not

17 This appears to be the path taken by Rodriguez (2017: 178–180) and Jones (2018) — understandably, since they mostly focus on *lipnē* and not on other expressions. In cognitive linguistics one typically tries not to assume many different homonymous lexemes,

grammaticalize: I still allow for semantic change in each prepositional expression (following common grammaticalization clines). I simply do not see them as entirely separate lexemes. This way, their meaning is composed of both "old" meaning common to most expressions built on *på̄nīm* and "new" meaning particular to specific expressions.

Let us now look at English *before* again. The claim I will put forward here is that this preposition had the same connotation of threatening influence in earlier stages of English, but that this connotation has been lost. The word therefore instantiates some parts of the semantic network given for Hebrew *pā̄nīm* in (5.16). In present-day English the meaning of *before* is spatial and temporal, although the spatial sense is already somewhat archaic when compared to *in front of* (OED: s.v. before). Etymologically, *before* is composed of the prefix *be-* 'about', inherited from Germanic, and the adverb *fore* 'in or of the front' (OED: s.v. before). This latter component is ultimately cognate with Latin *prō* and Ancient Greek πρό, both meaning 'for'; a lexeme, therefore, with a highly grammaticalized meaning for considerable time already. It is therefore hard to see whether *before* is ultimately related to the body part 'face'. Nevertheless, it is striking that in its spatial sense *before* refers particularly often to people (5.17), which might point to a relationship with a body part.

(5.17) a. *Before a crowded Dallas press conference, a pleased Perot unveiled new versions of the old props.*
(*Time International* 25 Jan. 1993, 18/1 in citation by OED: s.v. before, B.I.3a)

b. *An Italian proverb runs thus, "Who flatters me before, spatters me behind."*
(*Marlburian* 31 Jan. 1883, 3/1 in citation by OED: s.v. before, A.I.3b)

What we do know is that the element *fore* had, at some point, a generic meaning 'front' similar to *surface* and *facade*. This can be seen in fossilized expressions like *bring something to the fore*. Furthermore, if we go back to the seventeenth century, we find quite clear evidence for a connotation of threatening influence (5.18). Both of these examples describe fleeing from someone who is exercising a kind of threatening influence (cf. [5.11]). In (5.18a), the threatening connotation is also nicely paralleled by *darest not … look me in the face* in the next line. In present-day English, this connotation

nor many unrelated semantic functions, but rather to show how the various meanings of a lexeme interact with and depend on each other (e.g. Tyler & Evans 2003: 37–63).

can still be recognized in some idioms, like *bow before* (one bows, after all, only before entities with authority), but it is not productive anymore; in the OED, the last compelling example similar to these dates from 1931.[18]

(5.18) a. *Thou runn'st before me, shifting every place,*[19] *// And darest not stand, nor look me in the face.*
(Shakespeare 1600, *A Midsummer Night's Dream,* 423–424 in citation by OED: s.v. before, B.I.2b(b))

b. *This makes them flee before a shadow, and when none pursueth them, they run away from themselves.*
(Scott 1673, *Sermon before Honourable Military Company* in citation by OED: s.v. before, B.I.2b(b))

In the semantic-pragmatic network in (5.16) we saw that the connotation of threatening influence depends on the association with authority or dignity on the one hand and the generic spatial meaning 'front' on the other. The former could perhaps be provided from context in cases like (5.18), but we have seen that the latter meaning is by now only retained in fossilized expressions. Even the spatial meaning 'in front of' is becoming archaic, while the derived temporal one becomes more and more prevalent (OED: s.v. before). The prerequisites for the connotation of threatening influence, still attested in the seventeenth century, are not available anymore in present-day English. As a result, *before* has lost this connotation, and its use in translations of (5.1–5.2) has become misunderstood.

Finally, let us return to the root *ngp* 'defeat, inflict' seen in (5.1–5.2). Bearing in mind that English *before* used to have a connotation of threatening influence, the translations with 'before' make perfect sense: the defeated party is in the vicinity of the prepositional object but also threatened by it. The expression *X is defeated before Y* implies that *Y* has authority over *X*. Now that we know that *before*'s connotation of threatening influence was lost in English, we can understand why more recent translations and scholars went looking for an alternative. However, to be defeated *before* someone in the sense of (5.18) is very different from being defeated *by* someone: the first describes the *state* of being defeated when confronted with some other party that is more powerful, whereas the second describes the *event* of being defeated by that other party. The former does not imply any actual involve-

18 A 2001 example mentions retreating *before* a deity, but here the deity can hardly be seen as a causal agent; this may be a simple spatial usage in the sense 'in front of'.

19 I.e., changing your place continuously.

ment of the second party: in examples like (5.18ab) it is the mere presence of the complement of *before* that is relevant rather than any physical action on their part. The expression 'in the face of' therefore seems a better translation for the combination of *ngp* and *lipnē*. In the sections below I will address the question who the Agent of *ngp* is, if *lipnē* does not mark the Agent. I will argue that the Agent is Yahweh, who determines the outcome of battles.

Having made a positive case for a Locative interpretation of *lipnē* (i.e., 'in the face of' rather than the Agent marker interpretation 'by'), we should now also show that an Agent marker interpretation is unlikely regardless of any alternatives. Indeed, there are many reasons on various levels why it is very unlikely that *lipnē* would be an Agent marker. In the remainder of this chapter I discuss five reasons: one from narrative structure, two language-internal, one theoretical, and one typological.

5.2 Narrative structure: *ngp* 'defeat' in context

To properly appreciate the function of *lipnē* in contexts like (5.1–5.2) we must first understand the meaning of the verb *ngp* 'inflict, defeat'. We can do this by looking at the structure of the larger narrative in which it usually appears. Previously I have shown that the description of battles in the Hebrew Bible follows a specific prototypical scenario (Staps 2018: 169–171; cf. Van Wolde 2009: 59–60). One can distinguish seven stages, each of which is characterized by the use of specific verbal roots. Not every stage is necessarily explicitly mentioned in each episode, but the order in which they appear is fixed. The stages can be described as:

1. Gathering: the armies gather together (*ʾsp* 'gather'; *yʿd* 'meet up').
2. Motion: the armies move to the battleground (common verbs of motion, often *ʿlh* 'go up').
3. Preparation: the armies prepare themselves (*ḥnh* 'encamp'; *ʿrk* 'set in array').
4. Fighting: the physical conflict itself, that is, the act of fighting (*lḥm* 'fight').
5. Settlement: it becomes clear who wins (*lkd* 'seize' (of cities); *ntn* 'give' in the expression 'Yahweh gave X in Y's hand').
6. Physical conclusion: the final blow, determined by the previous stage (*nkh* 'strike').

7. Aftermath: various endings are found. The winner can take possession of the land of the opponent, they can destroy it, or the loser can flee and be chased.

Two examples with many of the stages are the following, with numbers in parentheses indicating the different stages:

(5.19) Josh. 10:34–35: וַיַּעֲבֹר יְהוֹשֻׁעַ וְכָל־יִשְׂרָאֵל עִמּוֹ מִלָּכִישׁ עֶגְלֹנָה וַיַּחֲנוּ עָלֶיהָ וַיִּלָּחֲמוּ עָלֶיהָ: וַיִּלְכְּדוּהָ בַּיּוֹם הַהוּא וַיַּכּוּהָ לְפִי־חָרֶב וְאֵת כָּל־הַנֶּפֶשׁ אֲשֶׁר־בָּהּ בַּיּוֹם הַהוּא הֶחֱרִים

way-y-aʿăḇōr-Ø *yəhōšuaʿ* *wə=ḵål* *yiśrāʾēl* *ʿimm-ō*
WAYQ-3M-pass_on-SG Joshua and=all.of Israel with-him

mil=lāḵīš *ʿɛḡlōn-āh* ***way-y-aḥăn-ū*** *ʿālɛ-hā*
from=Lachish Eglon-ALL **WAYQ-3M-encamp-PL** against-it

way-y-illāḥăm-ū *ʿālɛ-hā* ***way-y-ilkəḏ-ū-hā*** *b=ay=yōm*
WAYQ-3M-fight-PL against-it **WAYQ-3M-seize-PL**-it.OBJ on=the=day

ha=hū' ***way-y-akk-ū-hā*** *lə=pī* *ḥɛrɛḇ*
the=that **WAYQ-3M-strike-PL**-it.OBJ with=mouth.of sword

wə=ʾēṯ *ḵål* *han=nɛpɛš* *ʾăšɛr* *b-āh* *b=ay=yōm* *ha=hū'*
and=OBJ all.of the=being REL in-it on=the=day the=that

hɛḥĕrīm-Ø
destroy\PFV-3M.SG

'Then Joshua, and all of Israel with him, *passed on* (2) from Lachish to Eglon. They *encamped* (3) against it and *fought* (4) over it. They *seized* (5) it that day, and *struck* (6) it with the edge of the sword, and all the people that were in it he *destroyed* (7).'

(5.20) Josh. 10:29–30: וַיַּעֲבֹר יְהוֹשֻׁעַ וְכָל־יִשְׂרָאֵל עִמּוֹ מִמַּקֵּדָה לִבְנָה וַיִּלָּחֶם עִם־לִבְנָה: וַיִּתֵּן יְהוָה גַּם־אוֹתָהּ בְּיַד יִשְׂרָאֵל וְאֶת־מַלְכָּהּ וַיַּכֶּהָ לְפִי־חָרֶב וְאֶת־כָּל־הַנֶּפֶשׁ אֲשֶׁר־בָּהּ לֹא־הִשְׁאִיר בָּהּ שָׂרִיד

way-y-aʿăḇōr-Ø *yəhōšuaʿ* *wə=ḵål* *yiśrāʾēl* *ʿimm-ō*
WAYQ-3M-pass_on-SG Joshua and=all.of Israel with-him

mim=maqqēḏå *liḇnå* ***way-y-illåḥɛm-Ø*** *ʿim* *liḇnå*
from=Makkedah Libnah **WAYQ-3M-fight-SG** with Libnah

way-y-ittēn-Ø *yhwh* *gam* *ʾōṯ-åh* ***bə=yaḏ*** *yiśrå'ēl* *wə=ʾɛṯ*
WAYQ-3M-give-SG Yahweh also OBJ-it **in=hand.of** Israel and=OBJ

malk-åh ***way-y-akkɛ-Ø-hå*** *lə=pī* *ḥɛrɛḇ* *wə=ʾɛṯ*
king-its **WAYQ-3M-strike-SG**-it.OBJ with=mouth.of sword and=OBJ

kål	*han=nɛp̄ɛš*	*ʾăšɛr*	*b-ā̆h*	*lō'*	***hiš'īr-Ø***	*b-ā̆h*
all.of	the=being	REL	in-it	not	**remain\CAUS.PFV-3M.SG**	in-it

śā̆rīḏ
remnant

'Then Joshua, and all Israel with him, *passed on* (2) from Makkedah to Libnah, and they *fought* (4) with Libnah. And Yahweh *gave* it, too, *in the hand* (5) of Israel, as well as its king. And he[20] *struck* (6) it with the edge of the sword, and all the people that were in it: he *left no survivors* (7).'

The events described in war contexts of *ngp* fit the same seven-stage pattern. The data set underlying Staps (2018) only included episodes where the verb *lḥm* 'fight' is used, which covered only two instances of *ngp* (Deut. 1:42; 1 Sam. 4:10). Because in these two examples not all stages are explicit, they do not provide enough evidence to determine whether *ngp* belongs in stage 5 or stage 6. I then preliminarily placed them in stage 5 (Staps 2018: 177–181), based on the involvement of God with this root in the active voice (see section 5.3 below) and the meaning of the verb in general. However, an Agent marker interpretation would place *ngp* in stage 6, so I will make a better case for *ngp* as a stage 5 verb here.

First of all, it is clear that stage 5 and 6 are distinct stages, given that they frequently co-occur and are always in the same order (Staps 2018: 177–181). The physical conclusion is always performed by a human Agent, whereas the settlement can be determined by God (5.20). War in the Hebrew Bible is a religious event (cf. Walzer 1992, among others), so it is expected that Yahweh plays a pivotal role in this prototypical scenario. It is Yahweh who decides who wins; without him on one's side, one better does not go to battle at all (Deut. 1:42). However, while there is evidence for God's involvement in stage 5, the Agent in stage 6 is always human. God is seen as the divine orchestrator of the battle, determining the outcome, but he does not physically participate in it. He may play a different role in battles that are narrated differently, but not in the ones that adhere to this scenario.

The 24 or 25 occurrences of *ngp* in war contexts that were not yet considered in Staps (2018) also fit the pattern described above.[21] These examples

20 Based on a comparison with examples like (5.19) we must identify the subject of *strike* as Joshua, not Yahweh.

21 Lev. 26:17; Num. 14:42; Deut. 1:42; 28:7, 25; Jdg. 20:32, 35, 36, 39; 1 Sam. 4:2, 3, 10; 7:10; 2 Sam. 2:17; 10:15, 19; 18:7; 1 Kgs. 8:33; 14:12; Ps. 89:24 (unclear); 1 Chr. 19:16, 19; 2 Chr. 6:24; 13:15; 14:11; 20:22; 25:22. See section 5.3 below for a detailed breakdown.

allow us to be more specific as to the place of this root in the prototypical scenario. Concretely, there are instances where a stage 6 verb is present as well, confirming the classification of *ngp* as a stage 5 verb:

(5.21) Jdg. 20:35: וַיִּגֹּף יְהוָה | אֶת־בִּנְיָמִן לִפְנֵי יִשְׂרָאֵל וַיַּשְׁחִיתוּ בְנֵי יִשְׂרָאֵל בְּבִנְיָמִן בַּיּוֹם הַהוּא עֶשְׂרִים וַחֲמִשָּׁה אֶלֶף וּמֵאָה אִישׁ כָּל־אֵלֶּה שְׁלֵף חָרֶב:

way-y-iggōp-Ø	*yhwh*	*ʾɛṯ*	*binyå̄min*	*li=p̄n-ē*	*yiśrå̄ʾēl*
WAYQ-3M-defeat-SG	Yahweh	OBJ	Benjamin	to=face-of	Israel

way-y-ašḥīṯ-ū	*b̲ən-ē*	*yiśrå̄ʾēl*	*bə=b̲inyå̄min*
WAYQ-3M-destroy-PL	son-PL.of	Israel	against=Benjamin

b=ay=yōm	*ha=hūʾ*	*ʿɛśr-īm*	*wa=ḥămiššå̄*	*ʾɛlɛp̄*	*ū=mēʾå̄*	
on=the=day	the=that	ten-PL	and=five	thousand	and=hundred	
ʾīš	*kål*	*ʾēllɛ*	*šōlēp̄-Ø*	*ḥå̄rɛb̲*		
man	each.of	these	draw\PTCP-M.SG	sword		

'Then Yahweh *defeated* (5) Benjamin in the face of Israel. The Israelites *destroyed* (6) Benjamin that day: 25,100 men, each of them sword-drawing.'

(5.22) 1 Sam. 4:2: וַיַּעַרְכוּ פְלִשְׁתִּים לִקְרַאת יִשְׂרָאֵל וַתִּטֹּשׁ הַמִּלְחָמָה וַיִּנָּגֶף יִשְׂרָאֵל לִפְנֵי פְלִשְׁתִּים וַיַּכּוּ בַמַּעֲרָכָה בַּשָּׂדֶה כְּאַרְבַּעַת אֲלָפִים אִישׁ:

way-y-aʿar̲k-ū	*pəlištī-m*	*liqraʾṯ*	*yiśrå̄ʾēl*
WAYQ-3M-set_up-PL	Philistine-PL	opposite	Israel

wat-t-iṯṭōš-Ø	***ham=milḥå̄må̄***	***way-y-innå̄g̲ɛp̄-Ø***	
WAYQ-F-spread_out-3SG	**the=battle(F)**	**WAYQ-3M-defeat\MID-SG**	

yiśrå̄ʾēl	*li=p̄n-ē*	*p̄əlištī-m*	***way-y-akk-ū***
Israel	to=face-of	Philistine-PL	**WAYQ-3M-strike-PL**

b̲=am=maʿărå̄k̲å̄	*b=aś=śå̄d̲ɛ*	*kə=ʾarbaʿa-ṯ*	*ʾălå̄p̄-īm*	*ʾīš*	
in=the=battle_line	in=the=field	like=four-of	thousand-PL	man	

'Then the Philistines *arranged* (3) themselves opposite Israel. As *the battle* (4) spread out, Israel was *defeated* (5) in the face of the Philistines. They *struck* (6) around 4,000 men on the battlefield.'

These examples show that the root *ngp* is a stage 5 verb in both the active voice (qal template, [5.21]) and the non-active voice (niphal template, [5.22]).[22] In particular in the qal, where Yahweh is the subject, this is consis-

22 A number of recent studies have discussed the function of the niphal; see the excellent overview of the debate in Van Wolde (2019), to which may be added the response by Jones (2020). The point of discussion is to what extent the niphal expresses the passive and/or middle voice. Although we are concerned with *ngp* in the niphal, this discussion is only tangentially related to the matter at hand: my analysis that *lipnē* is not an Agent

tent with the observation that God plays a role in stage 5 but not stage 6. The fact that the typical order of events and the way the battle is narrated in general is the same regardless of voice suggests that the semantic content, and hence argument structure, of the verb is the same as well. In other words, in both the active and the non-active voice we must understand this root to have three arguments. In the active voice these can all be made overt, as in (5.21): the Agent as the subject, the Patient as the object, and a third argument marked by *lip̄nē*. In the non-active voice we expect the same three arguments in the underlying argument structure. It would be odd if *lip̄nē* marked the Agent in this situation, because this would (a) remove God from the argument structure when compared to the active voice and (b) make it impossible to express the third argument (which is expressed by *lip̄nē* in the active voice). Instead, we should understand an unexpressed Agent in the underlying argument structure. When *ngp* is used in the niphal, God's involvement is frequently made clear with other verbs in the direct context.[23] The simplest explanation is then that God is still the Agent in the events narrated with a niphal, but that the Agent is, as usual in Biblical Hebrew, not specified.[24] Therefore, *lip̄nē* does not mark the Agent.

The data of a different root, *knʿ* 'subdue', supports this. Like *ngp*, this verb, which uses the causative hiphil template for the active voice, is trivalent: God is the Agent, an army is the Patient, and a second army is marked by *lip̄nē* (5.23). The niphal of this root is used in the same way as the niphal of *ngp*: God is not mentioned, the Patient is the subject, and the second army is still marked by *lip̄nē* (5.24). Thus, by the same reasoning as above, *lip̄nē* cannot mark the Agent. Furthermore, this root also shows that the argument marked by *lip̄nē* cannot be an Intermediary Agent or Instrument, as

marker is compatible with both the position that the niphal expresses only the middle voice and the position that it expresses only the passive voice, as well as any position in between. For this reason I will agnostically refer to the niphal as a "non-active" voice.

23 Lev. 26:17 ('I will set my face against you'); Num. 14:42; Deut. 1:42 ('Yahweh is not among you'); Deut. 28:7, 25 ('Yahweh will allow … to happen'); 1 Sam. 7:10 ('Yahweh thundered loudly … against the Philistines and confused them'); 1 Kgs. 8:33; 2 Chr. 6:24 ('Israel sinned against you'); 2 Chr. 20:22 ('God wanted to give them in Joash' hand' in v. 20).

24 In Biblical Hebrew, "as a rule a proper passive form can be used only if the author of the action (*the agent*) is not named" (Joüon & Muraoka 2006: §132c, emphasis original). The normal way to express the Agent in Patient-oriented syntax is to use a relative clause. Thus for 'the innocent blood shed by Joab' we get 'the innocent blood which Joab shed' (1 Kgs. 2:31). For edge cases where prepositions might mark Agents, see Bicknell (1984: 43–51), but this is definitely not the default.

the hiphil would use a ditransitive construction for this (Joüon & Muraoka 2006: §125u).[25]

(5.23) Jdg. 4:23: וַיַּכְנַע אֱלֹהִים בַּיּוֹם הַהוּא אֶת יָבִין מֶלֶךְ־כְּנָעַן לִפְנֵי בְּנֵי יִשְׂרָאֵל׃

way-y-aḵnaʿ-Ø *ʾĕlōhīm* *b=ay=yōm* *ha=hūʾ* *ʾēṯ* *yāḇīn* *mɛlɛḵ*
WAYQ-3M-subdue-SG God on=the=day the=that OBJ Jabin king.of

kənāʿan ***li=pn-ē*** *bən-ē* *yiśrāʾēl*
Canaan **to=face-of** son-PL.of Israel

'On that day God subdued Jabin, the king of Canaan, *in the face of* the Israelites.'

(5.24) Jdg. 8:28: וַיִּכָּנַע מִדְיָן לִפְנֵי בְּנֵי יִשְׂרָאֵל

way-y-ikkāna-Ø *miḏyān* ***li=pn-ē*** *bən-ē* *yiśrāʾēl*
WAYQ-3M-subdue\MID-SG Midian **to=face-of** son-PL.of Israel

'Then Midian was subdued *in the face of* the Israelites.'

5.3 Syntactic-semantic restrictions: the arguments of *ngp* 'inflict, defeat'

Besides war contexts, the verb *ngp* 'inflict, defeat' is used in two other environments: in the sense of 'stumbling' and that of 'inflicting illness or injury'. A close look at the arguments it appears with in the different contexts yields two further arguments why the Agent marker interpretation of *lipnē* should be abolished. All 48 occurrences of the verb (not counting the infinitive absolute in Jdg. 20:39) are collected in table 5.1, along with the different arguments they have.

25 Of the studies I consulted, Sollamo (2003: 622–625) was the only one to take *lipnē* as an Agent marker with other verbs than *ngp* 'inflict, defeat' and *knʿ* 'subdue'. She sees the passive Agent as "an adjunct which expresses the originator or instrument of the action expressed by the passive predicate or participle and which in most cases can be converted into the subject of a corresponding active construction" (Sollamo 2003: 623). However, she does not take the semantic features of the arguments into account in her analysis. By her own definition, *lipnē* cannot be an Agent marker with *ngp* or *knʿ* because it marks a human argument while the subject of corresponding active constructions is always Yahweh. The other cases she considers are rare and doubtful. Most of them are best interpreted as a middle voice rather than a passive voice when one bases one's reading of the niphal verbs on more recent work like Van Wolde (2019); *lipnē* can be read in its default Locative sense and does not need to mark the Agent.

Meaning	Arguments			Occurrences
'Stumble'	Th: foot			2× qal; 1× hithpael[1]
'Inflict'	A: God	P: human		15 or 16× qal[2]
	A: God	P: land		1× qal[3]
	A: human	P: human		1× qal[4]
	A: animal	P: animal		1× qal[5]
'Defeat'	A: God	P: army	*lip̄nē*: army	4× qal[6]
	A: God	P: army		0 or 1× qal[7]
		P: army	*lip̄nē*: army	19× niphal[8]
		P: army		3× niphal[9]

1 In the qal: Ps. 91:12; Prov. 3:23. In the hithpael: Jer. 13:16.
2 Exod. 12:23 (2×), 27; 32:35; Josh. 24:5; 1 Sam. 25:38; 26:10; 2 Sam. 12:15; Isa. 19:22 (2×); Zech. 14:12, 18; Ps. 89:24 (which may also belong to the war context); 2 Chr. 13:20; 21:14, 18.
3 Exod. 7:27.
4 Exod. 21:22.
5 Exod. 21:35.
6 Jdg. 20:35; 1 Sam. 4:3; 2 Chr. 13:15; 14:11.
7 Ps. 89:24 (which may also belong to the injury context).
8 Lev. 26:17; Num. 14:42; Deut. 1:42; 28:7, 25; Jdg. 20:32, 39; 1 Sam. 4:2; 7:10; 2 Sam. 2:17; 10:15, 19; 18:7; 1 Kgs. 8:33; 14:12; 1 Chr. 19:16, 19; 2 Chr. 6:24; 25:22.
9 Jdg. 20:36; 1 Sam. 4:10; 2 Chr. 20:22.

Table 5.1 Contexts in which *ngp* occurs (A = Agent; P = Patient; Th = Theme).

In the meaning 'stumble', the root is intransitive and of little relevance to us here. However, the context of illness and injury is highly relevant, even though *lip̄nē* is not found in this environment. The exact semantic boundaries of this context are not exactly clear; it is best understood with a prototype model. The prototype is rather concrete and involves Yahweh 'inflicting' a disease upon humans. There is ample evidence for illness in a number of examples, such as the root *ʾnš* 'become ill' in (5.25).[26] When disease is mentioned explicitly the verb is always in the qal and the Agent is always Yahweh. Somewhat removed from this prototype, but still clearly related, are cases where God inflicts misfortune, but not necessarily disease (5.26). We also have two cases where humans and animals injure each other (Exod. 21:22, 35).

26 The involvement of illness is explicit in several other examples: 1 Sam. 25:38; Isa. 19:22 (2×); Zech. 14:12, 18; 2 Chr. 21:14, 18.

(5.25) 2 Sam. 12:15: וַיִּגֹּף יְהוָה אֶת־הַיֶּלֶד אֲשֶׁר יָלְדָה אֵשֶׁת־אוּרִיָּה לְדָוִד וַיֵּאָנַשׁ׃

way-y-iggōp-Ø *yhwh* *ʾɛṯ* *hay=yɛlɛḏ* *ʾăšɛr* *yāləḏ-āh*
WAYQ-3M-inflict-SG Yahweh OBJ the=child REL bear\PFV-3F.SG

ʾēšɛṯ *ʾūriyyā* *lə=ḏāwiḏ* *way-y-ēʾānaš-Ø*
wife.of Uriah to=David WAYQ-3M-become_ill-SG

'Then Yahweh *inflicted* the child which Uriah's wife bore to David, and it became ill.'

(5.26) Exod. 12:23: וְעָבַר יְהוָה לִנְגֹּף אֶת־מִצְרַיִם׃

wə-ʿāḇar-Ø *yhwh* *li=**ngōp*** *ʾɛṯ* *miṣrayim*
WQAT-pass_through-3M.SG Yahweh to=**inflict\INF** OBJ Egypt

'Yahweh will pass through to *strike/inflict something upon* the Egyptians.'

Observe that in almost every example God is the Agent of *ngp*. Furthermore, this is always the case for the war contexts in which *ngp* is in the active voice, like (5.27). Therefore, as already seen above, *ngp* is primarily a verb of divine intervention. If *lipnē* were an Agent marker, we would expect its complement to be God. This is however never the case: *lipnē* always marks an army, while armies are never the Agent in the active voice.

(5.27) 2 Chr. 14:11: וַיִּגֹּף יְהוָה אֶת־הַכּוּשִׁים לִפְנֵי אָסָא וְלִפְנֵי יְהוּדָה׃

way-y-iggōp-Ø *yhwh* *ʾɛṯ* *hak=kūšī-m* ***li=pn-ē*** *ʾāsāʾ*
WAYQ-3M-defeat-SG Yahweh OBJ the=Cushite-PL **to=face-of** Asa

*wə=**li=pn-ē*** *yəhūḏāh*
and=**to=face-of** Judah

'And Yahweh *defeated* the Cushites *in the face of* Asa and *in the face of* Judah.'

The second language-internal reason why *lipnē* should not be analyzed as an Agent marker is even clearer. In table 5.1 we see that *lipnē* is also used when *ngp* is in the active voice, as in (5.27). In the active voice, the Agent slot is taken by the subject. Therefore, with an Agent marker interpretation, *lipnē* must have another function in these four instances. A Locative interpretation does not suffer from this problem.[27]

27 Jones (2018: 225–226) solves this by reading *lipnē* as an Instrument marker in these cases: 'by/through/using'. But like an Agent marker reading, this interpretation as an Instrument marker can be ruled out on the basis of theoretical and typological arguments similar to the ones shown in the next sections. In short, it is unclear how *lipnē* would have developed from a spatial preposition to an Instrument marker, and we cannot explain why it would only be an Instrument marker in the context of *ngp* 'defeat'.

5.4 Grammaticalization and semantic restrictions

The fourth reason why I argue that *lipnē* is not an Agent marker stems from grammaticalization theory. Grammaticalization is the process as a result of which "particular items become more grammatical through time" (Hopper & Traugott 2003: 2). In our case, the particular item is the expression *lipnē*: literally this means 'to the face of', but over time it acquires the more grammatical senses of 'in front of' (spatial Locative) and 'before' (temporal Locative).[28] According to Rodriguez (2017: 180) this meaning has then further grammaticalized into Agent marking, and such a development is implicitly understood by Sollamo (2003) and Jones (2018) as well.

Grammaticalization, however, is not simply the acquiring of new, more grammatical, meanings. Grammaticalization theory provides a framework from which falsifiable predictions follow, against which we can check the likelihood of *lipnē* having developed an Agent marking function.[29] One of these predictions is that as a lexeme proceeds to acquire more grammatical meanings, it undergoes certain morphological changes that can be schematized by the cline in (5.28).

(5.28) content item > grammatical word > clitic > inflectional affix
(Hopper & Traugott 2003: 7)

Based on this cline it is clear that Agent marking is a highly grammatical function: in nominative-accusative languages with overt case marking, the Agent is marked with an inflectional affix in the active voice (the nominative case). Furthermore, many such languages mark the Agent in non-active voice with an inflectional affix as well, like the instrumental case in Russian. On the other hand, the spatial and temporal senses of *lipnē* ('before') are usually expressed with a preposition, i.e. a grammatical word (even if that preposition requires some case ending). These Locative senses are therefore less grammatical than Agent marking.

28 For our present purposes we can loosely define 'more grammatical' using the distinction between content words (*example, accept, green*) on the one hand and function words (*of, and, or, it, this*) on the other: the latter group is more grammatical than the former.

29 It is well-known that grammaticalization itself is unpredictable: we cannot predict the changes some construction may undergo. I use 'prediction' here in the sense of a practically testable implication of a hypothesis.

Furthermore, as words proceed down this cline, they lose semantic and pragmatic meanings: "as grammaticalized forms become increasingly syntacticized or morphologized they unquestionably cease over time to carry significant semantic or pragmatic meaning" (Hopper & Traugott 2003: 94). Thus, while the syntactic and morphological environment in which they occur may become more restricted, the semantic-pragmatic environment will become less restricted. Indeed, case-marking languages use the nominative case (which has reduced syntactic freedom and phonetic substance) for the subject regardless of the semantic or pragmatic context; this function is heavily embedded in the grammar.

We thus see that in grammaticalization two processes occur simultaneously: grammatical functions such as Agent marking are acquired, but this tends to go hand-in-hand with a reduction in semantic-pragmatic content ("bleaching"), syntactic freedom, and phonetic substance. It is unfortunate that this second aspect is not always taken into account by Hebraists. In the case of *lip̄nē*, there is no loss of pragmatic content (the connotation of threatening influence is still present), nor semantic bleaching (there are at most two verbs in combination with which *lip̄nē* would be an Agent marker). There is furthermore no evidence for reduction in syntactic freedom or phonetic substance. All in all, this is unlikely for a highly grammatical function like Agent marking. On the other hand, the Locative sense with its connotation of threatening influence that I propose is not only clearly less grammatical but also found in more semantic environments, as the example of *ʿāmaḏ lip̄nē* 'stand before' in the sense of 'serve' (5.10) has shown.

5.5 Typology: no arbitrary semantic shifts

Finally, when a word undergoes semantic shifts or grammaticalization, it does not acquire new meanings arbitrarily. These processes follow common paths fueled by cognitive processes. For example, *lip̄nē* came to mean 'in front of' because it describes the space close to or directed towards (*lə*) the face (*pn̊ānīm*). Since the processes through which expression meanings change are not particular to any language, we would expect similar cases in other languages — indeed, English *before* and *in front of* are parallels to *lip̄nē* in this sense. This spatial sense then developed into a temporal one ('before') on the basis of the widely attested TIME IS STATIONARY AND WE MOVE THROUGH IT metaphor (Lakoff & Johnson 1980: 43), in which events

are seen as locations. Again, English *before* has here undergone the same development as Biblical Hebrew *lip̄nē*.

However, it is not at all clear how one of the senses of *lip̄nē* would have developed into an Agent marking function, and Sollamo (2003), Rodriguez (2017), and Jones (2018) do not attempt to explain this. In default of a cognitive explanation, we can consider the a priori likelihood that *lip̄nē* developed an Agent marking function from its Locative meaning, by searching for expressions in other languages that may have undergone the same shift. The typological study of Palancar (2002) examines 148 languages with Agent markers (176 distinct markers), of which 87 languages have nominative-accusative alignment (106 passive Agent markers in total: Palancar 2002: 16–17). Palancar cross-references these markers with their other functions. In nominative-accusative languages, Agent markers commonly share the function of Agent marking with the function of marking Source (47%), Cause (33%), Locative (27%), and/or Instrument (23%) (Palancar 2002: 41–43).[30] At first sight, it may seem that the relatively high percentage of passive Agent markers with a Locative function (27%) suggests that *lip̄nē* may be an Agent marker as well. However, closer inspection shows that it is not common that a grammatical item has *only* Locative and passive Agent marking functions: in 14 out of the 18 times that Locative and passive Agent co-occur, other functions are involved as well, like Cause or Instrument. It seems that the Agent marker developed from such an intermediary category rather than directly from the Locative function (Palancar 2002: 206–207). There is no evidence for such an intermediary category in the case of *lip̄nē*.

Furthermore, those languages that do attest a direct development from Locative to passive Agent marker tend to be centred around Oceania, a pattern that is also found with ergative Agent markers (Palancar 2002: 259–261): Agent markers tend to develop directly from Locative markers with high frequency in Oceania, and the only two cases outside Oceania are of the Chukotko-Kamchatkan family, spoken in the far east of Russia. Direct development of Agent marking from Locative functions therefore seems to be an areal phenomenon. In contrast, the use of body parts as spatial prepositions and the relationship between 'face' and authority and dignity is ubiquitous. Hence it seems a priori highly unlikely for *lip̄nē* to have developed an Agent marking function, especially since no examples of Cause or Instrument marking have been found.

30 Palancar refers to the function of Source as "Ablative".

5.6 Related work

In the previous sections I have first made a positive case for a connotation of threatening influence for the word *pānīm* 'face' and the series of prepositional expressions derived from this noun. I have argued that this analysis is particularly useful to interpret occurrences of *lipnē* 'before' with the verb *ngp* 'defeat, inflict', and advanced five arguments why the recent interpretation of *lipnē* as an Agent marker 'by' in this context is incorrect. It is now time to discuss how this analysis relates to the work of the proponents of the Agent marker interpretation.

First of all, Sollamo defines the Agent as "an adjunct which expresses the originator or instrument of the action expressed by the passive predicate or participle and which in most cases can be converted into the subject of a corresponding active construction" (Sollamo 2003: 623). She notes that *lipnē* "usually means 'in the presence of, in front of, before'" and that this meaning is also found in passives, but goes on to mention "a few cases [in which] it is better to interpret *lipnē* as the preposition of an agent" (Sollamo 2003: 623). The argument is purely exegetical: it is based on the subjective question which interpretation is "better". In spite of this, Sollamo generally does not consider the wider context, such as the prototypical war scenario I described above, and thus relies too much on a translation like 'defeat' to decide whether a verb requires an Agent in the passive voice or not. For the combination with *ngp* 'defeat, inflict', she writes: "because *ngp lipnē* indicates the victor and not where the victory was won, it is simplest to regard the preposition *lipnē* as the preposition of the agent here" (Sollamo 2003: 624). She thus takes thematic roles as discrete: the argument is either Locative or Agent; since it cannot be Locative, it must be the Agent. In my analysis, the argument has indeed some Agent properties (such as sentience and the power to bring about an event or change of state), but not others (like actually bringing about that event or change of state, because a comparison with the active voice shows that Yahweh is the actual Agent).

It is also worth noting that Sollamo's definition of the passive Agent combines very broad semantic properties ("originator or instrument of the action") with syntactic ones ("adjunct"; "can be converted into the subject of a corresponding active construction"). This is broad to the point that it loses its explanatory power. Even if *lipnē* would mark the Agent with *ngp* according to this definition, it does not tell us whether the army would be

the "originator" of the action or the "instrument" of an action by some other originator. At the same time, we cannot narrow down the semantic aspects of the definition because they are tied to the syntactic properties. A syntactic subject need not be an Agent (in my stricter sense), as is evidenced by well-known examples like *The hammer broke the window*. Therefore, if the definition is to include syntactic aspects, it seems it necessarily loses explanatory power in semantics.

The next work to discuss the combination of *lip̄nē* with *ngp* is Rodriguez (2017).[31] He begins with the question whether *lip̄nē* is actually a "word", and argues that we should ask this question for each usage separately instead of the construction in general: we can "identify which utterances are composites of two words and which are utterances where the two words have become a frozen union" (Rodriguez 2017: 167–168). This is an enhancement of the position that *lip̄nē* is either always a "composite" (and hence structurally transparent) or always a "frozen union". I want to take this even further, arguing that there is no need to categorize usages as either "composite" or "frozen union", because even a frozen union with a grammaticalized meaning can retain meaning of the composite parts. Thus the grammaticalized Locative meaning of *lip̄nē* 'before' (in which we do not understand an actual face) still carries meaning of the noun *pn̄ām* (including the connotation of threatening influence). This allows me to distinguish between meaning that results from grammaticalization (Locative) and meaning that results from metaphorical extension (threatening influence). This way we do not need to assume a typologically unlikely grammaticalization path. Furthermore, this approach allows us to attribute the metaphorical extension to the noun *pn̄ām* rather than the prepositional expression *lip̄nē*. As a result, we have a more economical explanation for the same connotation with other prepositions and standalone *pn̄ām* (see [5.9–5.15]).

The last author to discuss here is Jones (2018). He starts out with a methodology much like the one I followed in section 5.3, investigating the number of arguments *ngp* 'defeat, inflict' requires and their semantic properties (Jones 2018: 220). However, he includes all usages of *ngp*, including

31 Hardy (2022: 169–175) also discusses *lip̄nē*, but not the combination with *ngp* specifically. In his analysis these cases are Locative, but it is not clear whether he understands there to be any metaphorical extension (such as threatening influence); given the immense scope of his work, there simply is no space to discuss such infrequent verb-preposition pairs.

the ones that occur in the context of illness and injury. As a result, he concludes that the Agent of *ngp* can be human (based on Exod. 21:22 in the illness and injury context), whereas such cases do not occur in the war context. There is good reason to separate these contexts: both the niphal of *ngp* and a third argument marked by *lipnē* only occur in the war context. That *lipnē* also occurs in the active voice is ignored in Jones's discussion of the arguments of *ngp*. He only discusses this briefly in his discussion of 1 Sam. 4:2 (Jones 2018: 225–227). This verse is crucial, because in the next verse *ngp* occurs in the qal with *lipnē*:

(5.29) 1 Sam. 4:2–3: וַיַּעַרְכוּ פְלִשְׁתִּים לִקְרַאת יִשְׂרָאֵל וַתִּטֹּשׁ הַמִּלְחָמָה וַיִּנָּגֶף יִשְׂרָאֵל לִפְנֵי פְלִשְׁתִּים וַיַּכּוּ בַמַּעֲרָכָה בַּשָּׂדֶה כְּאַרְבַּעַת אֲלָפִים אִישׁ׃

way-y-aʿarḵ-ū *pəlištī-m* *liqraʾṯ* *yiśrāʾēl*
WAYQ-3M-set_up-PL Philistine-PL opposite Israel

wat-t-iṭṭōš-Ø *ham=milḥāmā* ***way-y-innāgēp-Ø***
WAYQ-F-spread_out-3SG the=battle(F) **WAYQ-3M-defeat\MID-SG**

yiśrāʾēl ***li=pn-ē*** *pəlištī-m* *way-y-akkū*
Israel **to=face-of** Philistine-PL WAYQ-3M-strike-PL

b=am=maʿărāḵā *b=aś=śādɛ* *kə=ʾarbaʿa-ṯ* *ʾălāp-īm* *ʾīš*
in=the=battle_line in=the=field like=four-of thousand-PL man

'Then the Philistines arranged themselves opposite Israel. As the battle spread out, Israel *was defeated in the face of* the Philistines. They struck around 4,000 men on the battlefield.

וַיָּבֹא הָעָם אֶל־הַמַּחֲנֶה וַיֹּאמְרוּ זִקְנֵי יִשְׂרָאֵל לָמָּה נְגָפָנוּ יְהוָה הַיּוֹם לִפְנֵי פְלִשְׁתִּים

way-y-āḇō'-Ø *hā=ʿām* *ʾɛl* *ham=maḥănɛ* *way-y-ōʾmər-ū*
WAYQ-3M-come-SG the=people(M) to the=camp WAYQ-3M-say-PL

ziqn-ē *yiśrāʾēl* *lāmmā* ***nəgāp-Ø-ānū*** *yhwh* *hayyōm*
elder-PL.of Israel why **defeat\PFV-3M.SG**-us Yahweh today

li=pn-ē *pəlištī-m*
to=face-of Philistine-PL

When the people came to the camp, the elders of Israel said: "Why has Yahweh *defeated* us today *in the face of* the Philistines?"'

Jones admits the possibility that Yahweh is the "deleted agent in 4:2" (Jones 2018: 225), but prefers to take the Philistines as the Agent because they are involved in "striking around 4,000 men". He argues that the speech of the elders in 4:3 cannot be used to define the description of the event by the narrator in 4:2 (Jones 2018: 227). This is a valid observation, but by

pulling the two verses apart entirely Jones ignores the fact that 4:3 shows that the active voice of *ngp* often occurs with *lip̄nē*. For 4:3, Jones proposes to read *lip̄nē* as a marker of the Instrument ("Why did the Lord defeat us today *by* the Philistines?": 2018: 227, emphasis original). It is more economical, however, to attribute the same function to *lip̄nē* in all occurrences with *ngp*, both active and non-active. Jones does not accept Locative as a universal function, branding it as an "unhelpful and inappropriate translation" (Jones 2018: 217). However, he understands Locative in a very strict sense and does not consider possible metaphorical extensions. Furthermore, it is clear that he is working with the modern sense of English *before*, while I have shown that this translation is not as bad as one may think — only old-fashioned.

Finally, Jones observes that the niphal of *ngp* occurs almost exclusively with *lip̄nē*, implying that this suggests that it may mark the Agent: "It should at the very least be curious that the passive of a transitive action 'defeat' appears with a so-called locative *lip̄nē* in nearly every occurrence" (Jones 2018: 227). However, as I have shown above, the same is true for nearly every occurrence in the active voice (if we limit ourselves to the war context). Therefore, that *lip̄nē* is also found in the niphal is not so surprising, and an explanation should not be sought in its co-occurrence with this template. Jones is right to reject a strictly spatial Locative reading, but the solution can be found in the connotation of threatening influence instead of an entirely different function for *lip̄nē*.

5.7 Concluding remarks

In summary, translations and other scholars have proposed that *lip̄nē* is an Agent marker when used with *ngp* 'inflict, defeat' and perhaps a few other verbs. This is incorrect for five reasons: (a) the larger context in which this verb is used shows that the Agent of these events must be God, but *lip̄nē* marks humans; (b) the Agent of the verb *ngp* in the active voice is God, but again *lip̄nē* marks humans; (c) *lip̄nē* is also found with *ngp* in the active voice, where the Agent slot is already filled by the subject; (d) Agent marking is a highly grammaticalized function that should not be restricted to a specific semantic context; (e) it is cross-linguistically unlikely for a Locative preposition to directly develop an Agent marking function.

Instead, we have seen evidence from various languages that the body part 'face' is related to concepts of authority and dignity, which can give rise

to a connotation of threatening influence. This leads to the use of *lipnē* 'to the face of' with *ngp* 'inflict, defeat' in war contexts, as in (5.1): 'Israel was defeated *before* the Philistines' (ESV). Here, 'before' denotes the threatening influence of the Philistines, and a more suitable translation in present-day English would be 'in the face of'.

The exegetical implications of my proposal are clear. When *ngp* 'inflict, defeat' is used in the niphal template, it is the state of being defeated that is described — rather than the event.[32] Based on parallels with the active voice, it is clear that the implicit Agent effecting this state is God. He is left unexpressed because the niphal focuses on the state itself, not the event that caused it. The argument marked by *lipnē* has the same function as in the active voice: it marks the dominant party in that state; the party that has a threatening influence over the subject (e.g., Israel in [5.1]). The implication is that God has allowed Israel to be defeated in confrontation with the Philistines.

32 This matches with recent accounts of the middle voice expressed by the niphal. Testen (1998: 138) argues that the niphal ingressively describes the entering into the state described by the verb, whereas Van Wolde's 'resultative' category (2019: 467; 2021: 438) focuses primarily on the state itself rather than the entering into that state. The exact aspect is not important here, as long as the state of being defeated is in view.

Summary and implications 6

The preceding five chapters have described various ways in which abstract relations are expressed in spatial terms.

Chapter 1 introduced a new way to formalize polysemy in prepositions with both a spatial and a causal meaning. In particular, this approach can be used to derive the meaning of *by*-phrases in passives. It was shown that current analyses of such phrases are not suitable for languages with multiple Agent prepositions, such as French. The model introduced in chapter 1 correctly derives the difference in meaning between French *de* 'from' and *par* 'through', in spatial contexts, causal adjuncts, and passives. Unlike other accounts of *by*-phrases, this approach does not take the Agent preposition as accidentally homonymous with its spatial counterpart, but rather derives its meaning in a different semantic-syntactic context.

Chapter 2 applied the same model to causal adjuncts in Biblical Hebrew. Like French, Biblical Hebrew has (at least) two prepositions which have, among others, a spatial and a causal meaning: מִן *min* 'from (Source), out of (Reason, Cause)' and בְּ *bə* 'in (Locative), for (Reason), with (Instrument)'. It was shown that taking the spatial origin of these prepositions into account leads to a more precise description of the contrast between the two in causal environments where both are, in principle, felicitous.

Interestingly, while the spatial meanings of French *de* and Biblical Hebrew *min* are very similar, their causal meanings are not. In French, the greater distance expressed by *de* compared to *par* is interpreted as a decreased ability to affect the Patient. By contrast, in Biblical Hebrew, the fact that *min* points to the Origin is relevant as well: a causer at the origin of a causal model can affect the Patient to a greater extent. We thus see that the mapping of the causal domain onto the spatial domain can exploit different features: the mapping in Biblical Hebrew is primarily based on Origin

(though distance is to some extent involved as well); the mapping in French is only based on distance. This difference is not problematic. One may compare this with the two conceptualizations of time (Lakoff & Johnson 1980): one in which we are stationary and time moves (*there's a deadline **coming up***), and one in which we move through time (*the weeks that are **behind** us*). These two conceptualizations can coexist even within the same language, so there is no reason why perspectives on the causal chain could not be realized differently between languages.

The main contribution of part I consists in providing a way to describe spatial-causal polysemy in a compositional formal semantic analysis. Under this view, prepositions have an abstract meaning, expressing an *abstract spatial* relation between a Figure and a Ground. Whether a particular prepositional construction is felicitous depends on the availability of an interpretation for the abstract meaning of the preposition given the specific semantic type of the Figure and the Ground in context. Unlike traditional grammaticalization studies in which the physical meaning of a preposition is taken as primary, in this model I propose to derive the *physical* spatial meaning from an *abstract* spatial meaning. Therefore, on this view, the physical spatial meaning of a preposition has the same status as its causal (temporal, ...) meaning, because both are directly derived from the same abstract spatial meaning.

Crucially, the data surveyed in part I show that the abstract meaning of a preposition has a spatial component. Both the physical and the causal domain are cognitively represented with some kind of spatial structure. More work is needed to determine what kind of spatial structure is involved, exactly. It seems intuitive to think that physical space is cognitively represented using a three-dimensional axis system, but this does not need to be the same in other domains. In particular, causal dependencies were described in chapters 1 and 2 using causal models, which can be seen as directed graphs. In this representation there is no notion of x, y, and z coordinates, though the notion of distance can be defined and receives a causal interpretation. This allows us to draw a sharp line between spatial notions that *are* exploited in the causal domain, such as distance, and spatial notions that are *not* exploited in the causal domain (for example, angles can be computed in the physical but not the causal domain). In this way, more precise descriptions of polysemy involving spatial representations enable us to determine which spatial notions are more primitive than others.

In part II the focus was on the interpretation of abstract spatial notions in the left periphery, in which attitudes of Speaker and Addressee towards information content are marked. Chapter 3 mostly focused on the English demonstrative and complementizer *that*, with an excursus into Romance complementizers derived from Latin *quod*. It was argued that the complementizer *that* is not entirely bleached, but still preserves some semantic content. In particular, it still holds a [+distal] feature. While this [+distal] feature is typically interpreted as describing a property of physical space in the nominal domain, in the left periphery, it can be said to involve the Addressee, who is seen as "far" from the Speaker. A typical example of an effect of [+distal] marking in the left periphery is that information content marked as [+distal] is seen as a presupposition, that is, a proposition that the Speaker takes for granted to be shared with the Addressee. At the same time, a non-physical interpretation of [+distal] can also be used to explain the interpretation of non-exophoric demonstrative *that*. Chapter 3 presented a model in which the exact interpretation of [+distal] is derived from the syntactic position (in the CP or DP layer) and the type of referent (concrete, physical or abstract, conceptual). In doing so it provided a unified analysis of demonstrative and complementizer *that*.

The same model was then applied to the Biblical Hebrew clausal connective כִּי *kī* in chapter 4. This clausal connective can express a vast variety of functions, which scholarship has so far only been able to describe using a large number of historically related but synchronically unrelated lexical entries. Chapter 4 showed that a more economical account is possible using the notion of Addressee involvement — in particular, *kī* is marked for reference to Common Ground. This account only requires separate lexical entries for two lexicalized meanings, and can derive all other functions from context. In addition, this description is better able to explain contrasts between *kī* and related particles with a partial functional overlap.

Finally, part III considered the use of spatial terms to describe relations between people, with the example of Biblical Hebrew לִפְנֵי *lipnē* 'to the face of'. Like the constructions described in the first two parts, this case study shows very clearly that spatial terms can be used to describe relations in abstract domains. In this case, the noun *pānim* 'face' comes to describe an area in front of the face of the prepositional object. Although this area does not have clear boundaries, it is possible to express that something is 'to' or 'from'

someone's face. The face is an area in an abstract space, representing someone's authority and dignity, which can be challenged by placing something in that area.

Let me finish by pointing out three implications of these results. First of all, it is sometimes assumed in grammaticalization studies, implicitly or explicitly, that grammaticalized forms have lost all connection to earlier stages. The literature reviews in the chapters throughout this dissertation provide examples of this view: the Agent preposition *by* often receives a denotation entirely different from spatial *by* (section 1.4.2); the complementizer *that* is seen as semantically vacuous, apart from carrying a feature indicating that it introduces a tensed clause (section 3.1); and several authors explicitly state that the synchronic meaning of Biblical Hebrew *kī* cannot be related to the deictic function of the particle it derives from (section 4.1.2). By contrast, the analyses in the preceding chapters have shown that it is often more economical to assume that not all meaning bleached away, but that some meaning — specifically, *spatial* meaning — remains. This is in fact an expected feature of grammaticalization processes, but it is not always given enough thought. The formal models in chapters 1 and 3, together with their application in chapters 2 and 4, show that taking persistent spatial meaning into account actually leads to simpler and yet more precise synchronic descriptions.

Second, it is important to note that not all of the types of polysemy considered in this dissertation can accurately be described as metaphors. In the cognitive linguistic tradition, as well as in some studies on grammaticalization, syncretisms involving the spatial domain are sometimes described using Lakoffian metaphors like TIME IS SPACE. However, it is important to note that Lakoffian metaphors also cover things like AN ARGUMENT IS WAR, which are mostly reflected in lexical items (e.g., *your claims are **indefensible***; Lakoff & Johnson 1980). The polysemy studied in part III could perhaps be described this way. However, the prepositional polysemy discussed in part I and the spatial features persisting in complementizers discussed in part II show that spatial terms are well-suited to be used in functional items as well. This type of non-spatial extension is crucially different. A metaphor of the type AN ARGUMENT IS WAR is defeasible, and there is a point where the metaphor fails. It is a 'way of thinking about something', and in principle there are multiple, mutually compatible ways to think about the same do-

main, which may be selected by speakers depending on their culture and/or the specific context. By contrast, it appears that the grammaticalized spatial features in prepositions and complementizers are the only natural way to express causation and refer to information content without periphrasis. These non-spatial extensions appear to be more firmly anchored in our cognitive conceptualization. Though the exact distinction from metaphors remains a topic for further study, it is clear that such a distinction is needed.

The final point concerns the *kind of spatial notions* that are exploited to describe relations in non-spatial domains. Previous scholarship on prepositional polysemy has focused on the *relative location* between Figure and Ground (for example, the fact that Source prepositions often mark Causes: chapters 1 and 2). The studies in this dissertation show that the notion of *distance* is worth exploring in greater depth as well. This notion was relevant for non-spatial domains in all three parts. Causal distance is used to express properties of causal models used to describe causal relations (part I). In part II we saw that distance can also make a distinction between the Speaker ("near") and the Addressee ("far"). Finally, in part III, a small distance to a valuable entity (the face) received an emotive, threatening interpretation. The fact that distance appears in three entirely different domains is strong evidence for it being a primitive, but highly versatile cognitive concept that easily extends to other domains. More studies of this type are needed to obtain a complete overview of the spatial notions most frequently extended to other domains.

A primer in Biblical Hebrew A

Since three chapters in this dissertation discuss phenomena in Biblical Hebrew, this appendix provides the necessary background for readers unfamiliar with this language. For reasons of space only a very brief introduction can be given; reference works can be consulted for more information.

A.1 Orthography and phonology

Like many other Semitic languages, Biblical Hebrew was originally transmitted in a consonantal script. I take my examples from the Masoretic Text, a standard version of the Hebrew Bible which was vocalized by a group of scholars in the 7th to 9th centuries of the common era. Occasionally I will remark on alternative vocalizations or textual emendations, but in general I follow the Masoretic Text.

Transliteration in this dissertation follows the guidelines of the *Journal of Semitic Studies*. Tables A.1 and A.2 provide an overview of the consonant and vowel inventories, respectively. Some comments on these tables:

- Since the transliteration follows traditional conventions that may not be transparent to the general reader, phonetic values are provided as well. These follow the Tiberian tradition according to Edzard (2011). Some of the phonetic values are disputed, and in any case they reflect various sound changes after the demise of Biblical Hebrew as a living language. The phonetic values are listed to give the reader some support in pronunciation if desired; the precise phonetic rendering is irrelevant to us here.
- The consonants /*k m n p ṣ*/ have a different form at the end of a word, extending below the baseline or (in the case of /*m*/) closed.
- The consonants /*b g d k p t*/ are spirantized after vowels, indicated in the transliteration with a macron above or below the letter. In the

א	ʾ	[ʔ]	ח	*ḥ*	[ħ]	פ	*p*	[p]
ב	*b*	[b]	ט	*ṭ*	[ŧ]	פ ף	*p̄*	[f]
ב	*b̲*	[v]	י	*y*	[j]	צ ץ	*ṣ*	[s]
ג	*g*	[g]	כ ך	*k*	[k]	ק	*q*	[q]
ג	*ḡ*	[ɣ, ʁ]	כ ך	*k̲*	[x, χ]	ר	*r*	[ʀ, ʁ, ɾ]
ד	*d*	[d]	ל	*l*	[l]	שׂ	*ś*	[s]
ד	*d̲*	[ð]	מ ם	*m*	[m]	שׁ	*š*	[ʃ]
ה	*h*	[h]	נ ן	*n*	[n]	ת	*t*	[t]
ו	*w*	[w]	ס	*s*	[s]	ת	*t̲*	[θ]
ז	*z*	[z]	ע	ʿ	[ʕ]			

Table A.1 Transliteration and phonetic value of Biblical Hebrew consonants.

◌ָ	*a*	[a(:)]	◌ְ	*ə, Ø*	[a, Ø]
◌ָ	*å̄, å*	[ɔ(:)]	◌ֲ	*ă*	[a]
◌ֶ	*ɛ*	[ɛ(:)]	◌ֱ	*ĕ*	[ɛ]
◌ֵ	*ē*	[e(:)]	◌ֳ	*ă̆*	[ɔ]
◌ִ	*ī, i*	[i(:)]			
◌ֻ	*ū, u*	[u(:)]			
◌ֹ	*ō*	[o(:)]			

Table A.2 Transliteration and phonetic value of Biblical Hebrew vowels.

Hebrew script the non-spirantized pronunciation is indicated with a dot in the letter.

- The transliteration *yhwh* is used for the divine name יהוה 'Yahweh', which is traditionally not pronounced but replaced by *ʾăḏōnā́y* 'my lord' or *ʾĕlōhīm* 'God'.

A.2 Verbal morphology

Verbal forms are based on a triconsonantal root, which delineates a certain semantic field. Within this field, verbal templates consisting of a vowel pattern and, in some cases, affixes, derive a specific verbal meaning through specifying one of three different types of Aktionsart: simple, pluractional, and causative. The system of verbal templates (also called "stems" or "binyanim") is also used to specify active, passive, reflexive, and middle Voice. As shown in table A.3, not all Aktionsart-Voice combinations are productive in Biblical Hebrew anymore.[1] This table also indicates the glossing abbreviations used for these templates.

	Simple	**Pluractional**	**Causative**
Active	Qal (none)	Piel (PLURACT)	Hiphil (CAUS)
Passive	Passive qal (PASS)	Pual (PLURACT.PASS)	Hophal (CAUS.PASS)
Reflexive	n/a	Hithpael (PLURACT.REFL)	n/a
Middle	Niphal (MID)		

Table A.3 Names and glossing abbreviations for verbal templates.

In many cases the meaning of a root in a certain template has lexicalized. In such cases, I will not gloss Aktionsart and/or Voice but simply give an appropriate English translation.

1 The reflexive Voice of the simple and causative template dropped out of use. Though listed under simple Aktionsart, the niphal is not properly related to any type of Aktionsart and never had parallels in the pluractional or causative templates.

Finite forms are inflected for person, gender (masculine or feminine, except in the first person, which has common gender), and number (singular or plural). There are two main verbal conjugations: a suffix conjugation and a prefix conjugation. The suffix conjugation indicates person, gender, and number using suffixes and marks perfective/gnomic aspect and/or past tense. The prefix conjugation indicates person, gender, and number using both suffixes and prefixes, and marks imperfective aspect, non-past tense, and various modal nuances. Since these conjugations are traditionally called "perfect" and "imperfect", respectively, I use PFV and IPFV to gloss them. However, it should be kept in mind that they are not pure aspectual forms.

A dedicated conjugation is used to relate sequential events in the simple past in narrative texts. This conjugation, the *wayyiqṭol*, consists of a conjunction (*wa-*) followed by an old preterite (*yiqṭol*). The conjunction is semantically very light and can be translated as 'and', 'but', 'then', 'so', zero, or otherwise, depending on context. I gloss the conjunction in wayyiqṭol forms as WAYQ and assume a morpheme boundary rather than a clitic boundary after the conjunction.

A less frequent form is *wəqā̆ṭal*, of which the main function is to continue a previous modal form. This form consists of the conjunction *wə-* followed by the form of the suffix conjugation (*qā̆ṭal*), though in some forms the stress shifts. I gloss the conjunction in these forms as WQAT and again assume a morpheme boundary rather than a clitic boundary, to distinguish these forms from regular suffix conjugation forms following a conjunction.

In addition to these four main conjugations, Biblical Hebrew has an imperative for the second person (glossed IMP) and a jussive for the second/third person (glossed JUSS).[2]

There are two types of infinitives. The 'infinitive construct' is most similar to the infinitive or gerund in Indo-European languages; it is glossed as INF. The 'infinitive absolute' has various uses, the most notable of which is to strengthen a finite form of the same verb in the immediate context. This form is glossed INFABS.

Each verbal template has its own participle form, glossed PTCP. The qal template additionally has a passive participle, which is distinct in form but

2 The volitive of the first person exists but does not appear in the examples in this dissertation. The M.SG imperative, marked by *-Ø*, can be extended with a "paragogic" *-ā̆*, which has little semantic value; I treat it as a allomorph of the M.SG.

similar in meaning to the participle of the passive qal. I gloss the passive participle of the qal as PTCP.PASS (the participle of the passive qal, which does not appear in this dissertation, would be glossed PASS.PTCP).

When pronominal, objects are indicated with suffixes on the verb.[3] On the infinitive construct, a pronominal suffix may indicate both the subject and the object.

A.3 Nominal morphology

Nouns and adjectives are declined for gender (masculine or feminine) and number (singular, plural, or dual; the latter is of limited productivity). I do not gloss grammatical gender on nouns except where needed to clarify agreement features on verbs and adjectives.

Nominal forms can be in two 'states'. The absolute state is used in most environments. The construct state is used in possessive constructions to mark possessedness. Nominals in the construct state are immediately followed by their possessor. I gloss the construct state with '.of' or '-of', and do not mark the absolute state.

3 In some conjugations the object suffix can be preceded by an "energic" *n*, which has little semantic value; I do not gloss this morpheme.

Index of ancient sources

Bibliography

Al-Adaileh, Bilal A. & Renad Abbadi. 2012. The pragmatic implications of metonymical body-based idioms in Jordanian Arabic. *Argumentum* 8. 73–91.

Aejmelaeus, Anneli. 1986. Function and interpretation of כי in Biblical Hebrew. *Journal of Biblical Literature* 105(2). 193–209. DOI: 10.2307/3260389.

Angelopoulos, Nikos, Chris Collins & Arhonto Terzi. 2020. Greek and English passives, and the role of *by*-phrases. *Glossa: a journal of general linguistics* 5(1). 1–29. DOI: 10.5334/gjgl.1185.

Asher, Nicholas. 2011. *Lexical meaning in context: A web of words*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511793936.

Auer, Peter. 1998. Zwischen Parataxe und Hypotaxe: ‚Abhängige Hauptsätze' im gesprochenen und geschriebenen Deutsch [Between parataxis and hypotaxis: 'Dependent main clauses' in spoken and written German]. *Zeitschrift für germanistische Linguistik* 26(3). DOI: 10.1515/zfgl.1998.26.3.284.

Authier-Revuz, Jacqueline. 1972. Etude sur les formes passives du français [Study on the passive forms of French]. *Documentation et recherche en linguistique Allemande contemporain — Vincennes* 1. 1–145.

Axel-Tober, Katrin. 2017. The development of the declarative complementizer in German. *Language* 93(2). e29–e65. DOI: 10.1353/lan.2017.0030.

Bandstra, Barry Louis. 1982. *The syntax of particle ky in Biblical Hebrew and Ugaritic*. Yale University dissertation.

Bar-Asher Siegal, Elitzur A. 2013. *Introduction to the grammar of Jewish Babylonian Aramaic* (Lehrbücher Orientalischer Sprachen 3). Münster: Ugarit-Verlag.

Bar-Asher Siegal, Elitzur A. & Nora Boneh. 2020. Causation: From metaphysics to semantics and back. In Elitzur A. Bar-Asher Siegal & Nora Boneh (eds.), *Perspectives on causation: Selected papers from the Jerusalem 2017 workshop* (Jerusalem Studies in Philosophy and History of Science), 3–52. Cham: Springer. DOI: 10.1007/978-3-030-34308-8.

Barwise, Jon & John Perry. 1983. *Situations and attitudes*. Cambridge, MA: The MIT Press.

Bate, Danny L. In preparation. Where did *that* come from? The distribution and development of neutral finite complementizers in Indo-European.

Beal, Joan. 1988. Goodbye to all 'that'? The history and present behaviour of optional 'that'. In John Honey & Graham Nixon (eds.), *An historic tongue: Studies in English linguistics in memory of Barbara Strang*, 49–66. London: Routledge.

Beavers, John. 2011. On affectedness. *Natural Language & Linguistic Theory* 29(2). 335–370. DOI: 10.1007/s11049-011-9124-6.

Benigni, Antonella. 1999. The Biblical Hebrew particle כי from a discourse analysis perspective. *Zeitschrift für Althebraistik* 12(2). 126–145.

Biberauer, Theresa. 2017. Factors 2 and 3: A principled approach. *Cambridge Occasional Papers in Linguistics* 10(3). 38–65. DOI: 10.5565/rev/catjl.219.

Bicknell, Belinda Jean. 1984. *Passives in Biblical Hebrew*. University of Michigan dissertation.

Bierwisch, Manfred. 1999. How much space gets into language? In Paul Bloom, Merrill F. Garrett, Lynn Nadel & Mary A. Peterson (eds.), *Language and space*, 31–76. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/4107.003.0004.

Boas, Hans Christian. 2003. *A constructional approach to resultatives* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Bolinger, Dwight. 1972. *That's that* (Janua Linguarum 155). The Hague: Mouton.

Brasoveanu, Adrian & Donka F. Farkas. 2007. Say reports, assertion events and meaning dimensions. In Gabriela Alboiu, Andrei A. Avram, Larisa Avram & Daniela Isac (eds.), *Pitar Moş: A building with a view. Papers in honour of Alexandra Cornilescu*, 175–196. Bucharest: Editura Universităţii din Bucureşti.

Bruening, Benjamin. 2013. *By* phrases in passives and nominals. *Syntax* 16(1). 1–41. DOI: 10.1111/j.1467-9612.2012.00171.x.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26(5). 511–545. DOI: 10.1023/A:1025887707652.

Charnavel, Isabelle. 2017. Non-at-issueness of *since*-clauses. *Semantics and Linguistic Theory* 27. 43–58. DOI: 10.3765/salt.v27i0.4127.

Chen, Rong. 1990. English demonstratives: A case of semantic expansion. *Language Sciences* 12(2–3). 139–153. DOI: 10.1016/0388-0001(90)90009-6.

Cheshire, Jenny. 1996. That jacksprat: An interactional perspective on English *that*. *Journal of Pragmatics* 25(3). 369–393. DOI: 10.1016/0378-2166(95)00032-1.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Churchland, Paul M. 1986. Some reductive strategies in cognitive neurobiology. *Mind* 95(379). 279–309. DOI: 10.1093/mind/XCV.379.279.

Claassen, Walter T. 1983. Speaker-oriented functions of kî in Biblical Hebrew. *Journal of Northwest Semitic Languages* 11. 29–46.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620539.

Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805. DOI: 10.2307/414729.

Clédat, Leon. 1900. "De" et "par" après les verbes passifs ["De" and "par" after passive verbs]. *Revue de philologie française et de littérature* 14. 218–233.

Çokal, Derya, Patrick Sturt & Fernanda Ferreira. 2014. Deixis: *This* and *that* in written narrative discourse. *Discourse Processes* 51(3). 201–229. DOI: 10.1080/0163853X.2013.866484.

Colasanti, Valentina & Martina Wiltschko. 2019. Spatial and discourse deixis and the speech act structure of nominals. In *Proceedings of the 2019 annual conference of the Canadian Linguistics Association*, 1–14. http://hdl.handle.net/2262/100137 (28 August, 2023).

Collins, Chris. 2018. Is the passive *by*-phrase an adjunct? https://lingbuzz.net/lingbuzz/004032.

Conklin, Blane. 2011. *Oath formulas in Biblical Hebrew* (Linguistic Studies in Ancient West Semitic 5). Winona Lake, IN: Eisenbrauns.

Copley, Bridget & Heidi Harley. 2015. A force-theoretic framework for event structure. *Linguistics and Philosophy* 38(2). 103–158. DOI: 10.1007/s10988-015-9168-x.

Copley, Bridget & Heidi Harley. 2022. What would it take to tame the verbal Hydra? https://bcopley.com/wp-content/uploads/What-would-it-take-to-tame-the-verbal-Hydra-2022.pdf (28 August, 2023).

Corne, Chris. 1995. Nana k nana, nana k napa: The paratactic and hypotactic relative clauses of Réunion Creole. *Journal of Pidgin and Creole Languages* 10(1). 57–76. DOI: 10.1075/jpcl.10.1.03cor.

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. DOI: 10.1016/j.cognition.2008.06.010.

Croft, William. 2012. *Verbs: Aspect and causal structure*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199248582.001.0001.

Cruschina, Silvio & Eva-Maria Remberger. 2017. Before the complementizer: Adverb types and root clause modification. In Martin Hummel & Salvador Valera (eds.), *Adjective adverb interfaces in Romance* (Linguistics Today 242), 81–109. Amsterdam: John Benjamins. DOI: 10.1075/la.242.04cru.

Dahood, Mitchell. 1965. *Ugaritic-Hebrew philology. Marginal notes on recent publications* (Biblica et Orientalia 17). Rome: Pontifical Biblical Institute.

Dahood, Mitchell. 1966. *Psalms I, 1-50* (Anchor Bible Series 16). Garden City, NY: Anchor Bible.

Dahood, Mitchell. 1968. *Psalms II, 51-100* (Anchor Bible Series 17). Garden City, NY: Anchor Bible.

Dahood, Mitchell. 1970. *Psalms III, 101-150* (Anchor Bible Series 17A). Garden City, NY: Anchor Bible.

Dancygier, Barbara & Eve Sweetser. 2000. Constructions with *if*, *since*, and *because*: Causality, epistemic stance, and clause order. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause, condition, concession, contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 111–142. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043-006.

Davis, Anthony R. 2011. Thematic roles. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics*, vol. 1, 3 vols. (Handbooks of Linguistics and Communication Science 33), 399–420. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110226614.399.

De Vos, Mark. In preparation. An unexpected asymmetry between *because* and *for* causal adjuncts.

Dehé, Nicole & Anne Wichmann. 2010. Sentence-initial *I think* (*that*) and *I believe* (*that*): Prosodic evidence for uses as main clause, comment clause and discourse marker. *Studies in Language* 34(1). 36–74. DOI: 10.1075/sl.34.1.02deh.

Delsing, Lars-Olof. 2010. Exclamatives in Scandinavian. *Studia Linguistica* 64(1). 16–36. DOI: 10.1111/j.1467-9582.2010.01168.x.

Den Boon, Ton & Ruud Hendrickx (eds.). 2015–2017. *Dikke Van Dale Online*. https://www.vandale.nl.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization* (Typological Studies in Language 42). Amsterdam: John Benjamins. DOI: 10.1075/tsl.42.

Diessel, Holger & Michael Tomasello. 2001. The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12(2). 97–141. DOI: 10.1515/cogl.12.2.97.

Dirven, René. 1995. The construal of cause: The case of cause prepositions. In John R. Taylor & Robert E. MacLaury (eds.), *Language and the cognitive construal of the world* (Trends in Linguistics Studies and Monographs 82), 95–118. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110809305.95.

Dor, Daniel. 2005. Toward a semantic account of *that*-deletion in English. *Linguistics* 43(2). 345–382. DOI: 10.1515/ling.2005.43.2.345.

Doron, Edit. 2020. The causative component of psychological verbs. In Elitzur A. Bar-Asher Siegal & Nora Boneh (eds.), *Perspectives on causation: Selected papers from the Jerusalem 2017 workshop* (Jerusalem Studies in Philosophy and History of Science), 395–416. Cham: Springer. DOI: 10.1007/978-3-030-34308-8.

Dowty, David. 1989. On the semantic content of the notion of 'thematic role'. In Gennaro Chierchia, Barbara Hall Partee & Raymond Turner (eds.), *Properties,*

types and meaning. Vol. 2: *Semantic issues* (Studies in Linguistics and Philosophy 38–39). Dordrecht: Kluwer Academic. DOI: 10.1007/978-94-009-2723-0_3.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. DOI: 10.1353/lan.1991.0021.

Dowty, David. 2003. The dual analysis of adjuncts/complements in Categorial Grammar. In Ewald Lang, Claudia Maienborn & Cathrine Fabricius-Hansen (eds.), *Modifying adjuncts* (Interface Explorations 4), 33–66. De Gruyter Mouton. DOI: 10.1515/9783110894646.33.

Edzard, Lutz. 2011. Biblical Hebrew. In Stefan Weninger (ed.), *The Semitic languages* (Handbooks of Linguistics and Communication Science 36), 480–514. De Gruyter Mouton. DOI: 10.1515/9783110251586.480.

Elsness, Johan. 1984. *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies* 65(6). 519–533. DOI: 10.1080/00138388408598357.

Etxepare, Ricardo. 1997. *The grammatical representation of speech events*. University of Maryland dissertation.

Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120(3). 604–627. DOI: 10.1016/j.lingua.2008.07.009.

Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.

Farkas, Donka F. & Kim B. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27(1). 81–118. DOI: 10.1093/jos/ffp010.

Fillmore, Charles. 1968. The case for case. In Emmon Bach & Robert T. Harms (eds.), *Universals in linguistic theory*, 1–90. New York, NY: Holt, Rinehart & Winston.

Foley, William A. & Robert D. Van Valin. 1984. *Functional syntax and universal grammar* (Cambridge Studies in Linguistics 38). Cambridge: Cambridge University Press.

Folli, Raffaella & Heidi Harley. 2008. Teleology and animacy in external arguments. *Lingua* 118(2). 190–202. DOI: 10.1016/j.lingua.2007.02.004.

Follingstad, Carl M. 2001. *Deictic viewpoint in Biblical Hebrew text: A syntagmatic and paradigmatic analysis of the particle* כי (*kî*) (Journal of Translation and Textlinguistics). Dallas, TX: SIL International.

Gaatone, David. 1998. *Le passif en français* [The French passive] (Champs Linguistiques). Paris: Duculot.

Gesenius, Wilhelm. 1910. *Gesenius' Hebrew grammar*. Emil Kautzsch & Arthur Ernest Cowley (eds.). Oxford: Clarendon Press.

Ginzburg, Jonathan. 1996. Dynamics and the semantics of dialogue. In Jerry Seligman & Dag Westerståhl (eds.), *Language, logic, and computation*, vol. 1 (CSLI Lecture Notes), 221–237. Stanford, CA: CSLI.

Gordis, Robert. 1943. The asseverative kaph in Ugaritic and Hebrew. *Journal of the American Oriental Society* 63(2). 176–178. DOI: 10.2307/594125.

Gougenheim, Georges. 1938. *Système grammatical de la langue française* [Grammatical system of the French language]. Paris: d'Artrey.

Grice, Paul. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Gross, Walter. 1991. Satzfolge, Satzteilfolge und Satzart als Kriterien der Subkategorisierung Hebräischer Konjunktionalsätze, am Beispiel der כי-Sätze untersucht [Sentence order, clause order, and sentence type as criteria for the subcategorization of Hebrew conjunctional clauses, studied using the example of כי-clauses]. In Walter Gross, Hubert Irsigler & Theodor Seidl (eds.), *Text, Methode und Grammatik: Wolfgang Richter zum 65. Geburtstag* [Text, method, and grammar: Wolfgang Richter on his 65th birthday], 97–119. St. Ottilien: EOS Verlag.

Grosz, Barbara & Yael Ziv. 1998. Centering, global focus, and right dislocation. In Marilyn A. Walker, Aravind K. Joshi & Ellen F. Prince (eds.), *Centering theory in discourse*, 293–308. Oxford: Clarendon Press.

Gruber, Jeffrey. 1965. *Studies in lexical relations*. Indiana University dissertation.

Gutiérrez-Rexach, Javier. 2001. Spanish exclamatives and the interpretation of the left periphery. In Yves D'hulst, Johan Rooryck & Jan Schroten (eds.), *Romance languages and linguistic theory 1999* (Current Issues in Linguistic Theory 221), 167–194. Amsterdam: John Benjamins. DOI: 10.1075/cilt.221.07gut.

Gzella, Holger. 2008. Zum periphrastischen Infinitiv in Genesis viii 5 [On the periphrastic infinitive in Genesis 8:5]. *Vetus Testamentum* 58(4). 469–479. DOI: 10.1163/156853308X312717.

Haber, Esther. 2009. Significance of the interchanging of the mono-consonantal prepositions in Biblical Hebrew: Rabbinic and comparative Semitic evidence. *Maarav* 16(1). 29–37. DOI: 10.1086/MAR200916103.

Hallman, Peter. 2021. On passive and perfect participles. In Kleanthes K. Grohmann, Akemi Matsuya & Eva-Maria Remberger (eds.), *Passives cross-linguistically: Theoretical and experimental approaches* (Empirical Approaches to Linguistic Theory 17), 64–97. Leiden: Brill. DOI: 10.1163/9789004433427.

Halpern, Joseph Y. & Judea Pearl. 2005. Causes and explanations: A structural-model approach. Part I: Causes. *The British Journal for the Philosophy of Science* 56(4). 843–887. DOI: 10.1093/bjps/axi147.

Hardy, Humphrey Hill II. 2022. *The development of Biblical Hebrew prepositions* (Ancient Near East Monographs 28). Atlanta, GA: SBL Press.

Harley, Heidi. 2011. Thematic roles. In Patrick Colm Hogan (ed.), *The Cambridge encyclopedia of the language sciences*, 861–862. Cambridge: Cambridge University Press.

Hasselbach, Rebecca. 2007. Demonstratives in Semitic. *Journal of the American Oriental Society* 127(1). 1–27. JSTOR: 20297208.

Hernanz, M. Lluïsa. 2007. From polarity to modality. Some (a)symmetries between *bien* and *sí* in Spanish. In Luis Eguren & Olga Fernández Soriano (eds.), *Coreference, modality, and focus: Studies on the syntax-semantics interface* (Linguistics Today 111), 133–169. Amsterdam: John Benjamins. DOI: 10.1075/la.111.08her.

Hill, Virginia. 2012. A main clause complementizer. In Lobke Aelbrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena: New horizons* (Linguistics Today 190), 279–296. Amsterdam: John Benjamins. DOI: 10.1075/la.190.12hill.

Hoffmann, Sebastian. 2005. *Grammaticalization and English complex prepositions: A corpus-based study* (Routledge Advances in Corpus Linguistics 7). London: Routledge.

Holmstedt, Robert D. 2010. *Ruth: A handbook on the Hebrew text* (Baylor Handbook on the Hebrew Bible). Waco, TX: Baylor University Press.

Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497. JSTOR: 4177789.

Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*. Vol. 1: *Focus on theoretical and methodological issues* (Typological Studies in Language 19), 17–35. Amsterdam: John Benjamins. DOI: 10.1075/tsl.19.1.04hop.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299. DOI: 10.2307/413757.

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139165525.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. DOI: 10.1017/9781316423530.

Ingason, Anton Karl, Iris Edda Nowenstein & Einar Freyr Sigurðsson. 2016. The Voice-adjunction theory of agentive 'by'-phrases and the Icelandic impersonal passive. *Working Papers in Scandinavian Syntax* 97. 40–56.

Jackendoff, Ray. 1996. The architecture of the linguistic-spatial interface. In Paul Bloom, Mary A. Peterson, Lynn Nadel & Merrill F. Garrett (eds.), *Language and*

space, 1–30. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/4107.003.0003.

Jenni, Ernst. 1992. *Die hebräischen Präpositionen* [The Hebrew prepositions]. Vol. 1: *Die Präposition Beth* [The preposition beth]. 3 vols. Stuttgart: Kohlhammer.

Jenni, Ernst. 2000. *Die hebräischen Präpositionen* [The Hebrew prepositions]. Vol. 3: *Die Präposition Lamed* [The preposition lamed]. 3 vols. Stuttgart: Kohlhammer.

Jones, Ethan C. 2018. More than "before": Semantics and syntax of לפני. In Ethan C. Jones (ed.), *The unfolding of your words gives light: Studies on Biblical Hebrew in honor of George L. Klein*, 217–232. University Park, PA: Eisenbrauns.

Jones, Ethan C. 2020. Middle and passive voice: Semantic distinctions of the niphal in Biblical Hebrew. *Zeitschrift für die alttestamentliche Wissenschaft* 132(3). 427–448. DOI: 10.1515/zaw-2020-3004.

Jouitteau, Mélanie. 2004. Gestures as expletives: Multichannel syntax. In Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez & Benjamin Schmeiser (eds.), *WCCFL 23 proceedings. Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 422–435. Somerville, MA: Cascadilla Press. https://shs.hal.science/halshs-00605736 (28 August, 2023).

Joüon, Paul & Takamitsu Muraoka. 2006. *A grammar of Biblical Hebrew*. 2nd edn. Rome: Gregorian & Biblical Press.

Jungbluth, Konstanze. 2003. Deictics in the conversational dyad: Findings in Spanish and some cross-linguistic outlines. In Friedrich Lenz (ed.), *Deictic conceptualisation of space, time and person* (Pragmatics & Beyond New Series 112), 13–40. Amsterdam: John Benjamins. DOI: 10.1075/pbns.112.04jun.

Kadmon, Nirit. 2001. *Formal pragmatics: Semantics, pragmatics, presupposition, and focus*. Malden, MA: Blackwell.

Kajzer-Wietrzny, Marta. 2018. Interpretese vs. non-native language use: The case of optional *that*. In Mariachiara Russo, Claudio Bendazzoli & Bart Defrancq (eds.), *Making way in corpus-based interpreting studies* (New Frontiers in Translation Studies), 97–113. Singapore: Springer. DOI: 10.1007/978-981-10-6199-8_6.

Kaltenböck, Gunther. 2006. '*... That* is the question': Complementizer omission in extraposed *that*-clauses. *English Language and Linguistics* 10(2). 371–396. DOI: 10.1017/S1360674306001961.

Kaltenböck, Gunther. 2009. Initial *I think*: Main or comment clause? *Discourse and Interaction* 2(1). 49–70.

Katz, Graham. 2003. Event arguments, adverb selection, and the Stative Adverb Gap. In Ewald Lang, Claudia Maienborn & Cathrine Fabricius-Hansen (eds.), *Modifying adjuncts* (Interface Explorations 4), 455–474. Berlin: Mouton De Gruyter. DOI: 10.1515/9783110894646.455.

Kayne, Richard S. 2014. Why isn't *this* a complementizer? In Peter Svenonius (ed.), *Functional structure from top to toe*, vol. 9 (The Cartography of Syntactic Structures), 188–231. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199740390.003.0007.

Kimron, Elisha. 1980. שמות רוחות השמים במקורותינו הקדומים [The names of the cardinal directions in our ancient sources]. *Beit Mikra* 26(1). 41–47. JSTOR: 23503942.

Kirsner, Robert S. 1979. Deixis in discourse: An exploratory quantitative study of the Modern Dutch demonstrative adjectives. In Talmy Givón (ed.), *Discourse and syntax* (Syntax and Semantics 12), 355–375. Leiden: Brill. DOI: 10.1163/9789004368897_016.

Kocher, Anna. 2022. *Complementizers on edge: On the boundaries between syntax and pragmatics in Ibero-Romance* (Open Romance Linguistics 1). Berlin: Language Science Press. DOI: 10.5281/zenodo.6907848.

Krapova, Iliyana. 2010. Bulgarian relative and factive clauses with an invariant complementizer. *Lingua* 120(5). 1240–1272. DOI: 10.1016/j.lingua.2009.08.002.

Kraska-Szlenk, Iwona. 2014. Semantic extensions of body part terms: Common patterns and their interpretation. *Language Sciences* 44. 15–39. DOI: 10.1016/j.langsci.2014.02.002.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon* (Studies in Natural Language and Linguistic Theory 33), 109–137. Dordrecht: Springer. DOI: 10.1007/978-94-015-8617-7_5.

Kratzer, Angelika. 2000. Building statives. In Lisa J. Conathan, Jeff Good, Darya Kavitskaya, Alyssa B. Wulf & Alan C. L. Yu (eds.), *Proceedings of the twenty-sixth annual meeting of the Berkeley Linguistics Society*, 385–399. Berkeley, CA. DOI: 10.3765/bls.v26i1.1131.

Kratzer, Angelika, Luigi Rizzi, Roberta Pires de Oliveira, Ina Emmel & Monica Deitos Stedile. 2020. Informal formal conversation on syntax-semantics. In Roberta Pires de Oliveira, Ina Emmel & Sandra Quarezemin (eds.), *Brazilian Portuguese, syntax and semantics. 20 Years of Núcleo de Estudos Gramaticais*, 7–29. John Benjamins. DOI: 10.1075/la.260.01kra.

Krüger, Thomas. 2004. *Qoheleth: A commentary* (Hermeneia: A Critical and Historical Commentary on the Bible). Minneapolis, MN: Fortress Press.

Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World lexicon of grammaticalization*. 2nd edn. Cambridge: Cambridge University Press. DOI: 10.1017/9781316479704.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lakoff, Robin. 1974. Remarks on 'this' and 'that'. In Michael W. La Galy, Robert A. Fox & Anthony Bruck (eds.), *Papers from the tenth regional meeting of the Chicago Linguistics Society*, 345–356. Chicago, IL: Chicago Linguistic Society.

Landman, Fred. 2000. *Events and plurality* (Studies in Linguistics and Philosophy 76). Dordrecht: Springer. DOI: 10.1007/978-94-011-4359-2.

Lasnik, Howard & Mamoru Saito. 1991. On the subject of infinitives. In Lise M. Dobrin, Lynn Nichols & Rosa M. Rodriguez (eds.), *Papers from the 27th regional meeting of the Chicago Linguistic Society. Part one: The general session*, 324–343. Chicago, IL: Chicago Linguistic Society.

Ledgeway, Adam. 2011. Subject licensing in CP: The Neapolitan double-subject construction. In Paola Benincà & Nicola Munaro (eds.), *Mapping the left periphery: The cartography of syntactic structures*, vol. 5 (The Cartography of Syntactic Structures), 257–296. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199740376.001.0001.

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese* (Linguistic Inquiry Monographs 69). Cambridge, MA: The MIT Press. DOI: 10 . 7551 / mitpress / 9780262028141.001.0001.

Levinsohn, Stephen H. 2011. רַק and אַךְ: Limiting and countering. *Hebrew Studies* 52. 83–105. DOI: 10.1353/hbr.2011.0023.

Lewis, David K. 1969. *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.

Lewis, David K. 1973a. Causation. *The Journal of Philosophy* 70(17). 556–567. DOI: 10.2307/2025310.

Lewis, David K. 1973b. *Counterfactuals*. Malden, MA: Blackwell.

Lipiński, Edward. 2001. *Semitic languages: Outline of a comparative grammar*. 2nd edn. (Orientalia Lovaniensia Analecta 80). Leuven: Peeters.

Littlemore, Jeannette. 2019. *Metaphors in the mind: Sources of variation in embodied metaphor*. Cambridge: Cambridge University Press. DOI: 10.1017/9781108241441.

Locatell, Christian S. 2017. *Grammatical polysemy in the Hebrew Bible: A cognitive linguistic approach to* כִּי. University of Stellenbosch dissertation.

Locatell, Christian S. 2020. Temporal conjunctions and their semantic extensions: The case of כִּי in Biblical Hebrew. *Journal of Semitic Studies* 65(1). 93–115. DOI: 10.1093/jss/fgz040.

Mackie, John L. 1965. Causes and conditions. *American Philosophical Quarterly* 2(4). 245–264. JSTOR: 20009173.

Maes, Alfons, Emiel Krahmer & David Peeters. 2022. Explaining variance in writers' use of demonstratives: A corpus study demonstrating the importance of discourse genre. *Glossa: a journal of general linguistics* 7(1). DOI: 10.16995/glossa.5826.

Maienborn, Claudia & Johanna Herdtfelder. 2017. Eventive versus stative causation: The case of German causal *von*-modifiers. *Linguistics and Philosophy* 40(3). 279–320. DOI: 10.1007/s10988-016-9201-8.

Malle, Bertram F. & Joshua Knobe. 1997. The folk concept of intentionality. *Journal of Experimental Social Psychology* 33(2). 101–121. DOI: 10.1006/jesp.1996.1314.

Mannati, Marina. 1970. Sur le sens de *min* en Ps iv 8 [On the meaning of *min* in Ps. 4:8]. *Vetus Testamentum* 20(3). 361–366. DOI: 10.1163/156853370X00390.

Martin, Fabienne. 2005. Genitives of substance and locative verbs in French. In Jiyung Kim, Yury A. Lander & Barbara H. Partee (eds.), *Possessives and beyond: Semantics and syntax* (University of Massachusetts Occasional Papers in Linguistics 29), 411–426. Amherst, MA: GLSA Publications.

Martin, Fabienne & Florian Schäfer. 2013. On the argument structure of verbs with mono- and bi-eventive uses. In Shayne Sloggett & Stefan Keine (eds.), *Proceedings of the 42nd meeting of the North East Linguistic Society*, 297–308. Amherst, MA: Graduate Linguistic Student Association.

McCarthy, Dennis J. 1980. The uses of *w^{e}hinnēh* in Biblical Hebrew. *Biblica* 61(3). 330–342. JSTOR: 42706879.

Meyer, Esias E. 2001. The particle כִּי, a mere conjunction or something more? *Journal of Northwest Semitic Languages* 27(1). 39–62.

Michaelis, Laura. 2001. Exclamative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*, vol. 2, 2 vols. (Handbooks of Linguistics and Communication Science 20), 1038–1058. Berlin: De Gruyter. DOI: 10.1515/9783110194265-016.

Michotte, Albert. 1946. *La perception de la causalité* [The perception of causality] (Etudes de Psychologie 6). Louvain: Editions de l'Institut Supérieur de Philosophie.

Miller, Cynthia L. 2003. *The representation of speech in Biblical Hebrew narrative: A linguistic analysis* (Harvard Semitic Monographs 55). Boston: Brill.

Moltmann, Friederike. 2007. Events, tropes, and truthmaking. *Philosophical Studies* 134(3). 363–403. DOI: 10.1007/s11098-005-0898-4.

Moody, Raymond. 1972. Portuguese prepositions: Some semantic categories. *Luso-Brazilian Review* 9(1). 36–71. JSTOR: 3512838.

Morrill, Glyn V. 1994. *Type logical grammar: Categorial logic of signs*. Dordrecht: Springer. DOI: 10.1007/978-94-011-1042-6.

Moshavi, Adina. 2009. Two types of argumentation involving rhetorical questions in Biblical Hebrew dialogue. *Biblica* 90(1). 32–46. JSTOR: 42614874.

Muilenburg, James. 1961. The linguistic and rhetorical usages of the particle כי in the Old Testament. *Hebrew Union College Annual* 32. 135–160. JSTOR: 23524614.

Muraoka, Takamitsu. 1985. *Emphatic words and structures in Biblical Hebrew*. Jerusalem: The Magnes Press.

Næss, Åshild. 2007. *Prototypical transitivity* (Typological Studies in Language 72). Amsterdam: John Benjamins. DOI: 10.1075/tsl.72.

Neeleman, Ad & Hans van de Koot. 2012. The linguistic expression of causation. In Martin Everaert, Marijana Marelj & Tal Siloni (eds.), *The Theta System: Argument structure at the interface* (Oxford Linguistics), 20–51. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199602513.003.0002.

OED Online. 2021. Oxford University Press. https://www.oed.com/.

Oeming, Manfred. 2006. שָׁכַר. In G. Johannes Botterweck, Helmer Ringgren & Heinz-Josef Fabry (eds.), *Theological dictionary of the Old Testament*, vol. 15, 16 vols., 1–5. Grand Rapids, MI: William B. Eerdmans.

Palancar, Enrique L. 2002. *The origin of agent markers* (Studia Typologica 5). Berlin: Akademie Verlag. DOI: 10.1524/9783050081410.

Park, Grace J. 2016. Stand-alone nominalizations formed with *ʾăšer* and *kî* in Biblical Hebrew. *Journal of Semitic Studies* 61(1). 41–65. DOI: 10.1093/jss/fgv030.

Partee, Barbara H. 1997. Genitives: A case study. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of logic & language*, 464–473. Amsterdam: Elsevier.

Pearl, Judea. 2000. *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press. DOI: 10.1017/S0266466603004109.

Peeters, David, Emiel Krahmer & Alfons Maes. 2021. A conceptual framework for the study of demonstrative reference. *Psychonomic Bulletin & Review* 28(2). 409–433. DOI: 10.3758/s13423-020-01822-8.

Pérez, Carlos Muñoz & Matías Verdecchia. 2022. Information-based "island effects" in Spanish clausal doubling. Paper presented at the 45th Generative Linguistics in the Old World colloquium (GLOW 45), Queen Mary University of London, 27–29 April.

Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: The MIT Press.

Poschmann, Claudia. 2008. All declarative questions are attributive? *Belgian Journal of Linguistics* 22. 247–269. DOI: 10.1075/bjl.22.12pos.

Primus, Beatrice. 2016. Participant roles. In Nick Riemer (ed.), *The Routledge handbook of semantics* (Routledge Handbooks in Linguistics), 403–418. London: Routledge.

Pullum, Geoffrey K. 2006. Phrasal prepositions in a civil tone. In Mark Liberman & Geoffrey K. Pullum (eds.), *Far from the madding gerund and other dispatches from Language Log*, 41–45. Wilsonville, OR: William, James & Co.

Pylkkänen, Mariliina. 1999. On stativity and causation. In Carol L. Tenny & James Pustejovsky (eds.), *Events as grammatical objects: The converging perspectives of lexical semantics and syntax*, 417–445. Stanford, CA: CSLI Publications.

Pylkkänen, Liina. 2008. *Introducing arguments* (Linguistic Inquiry Monograph 49). Cambridge, MA: MIT Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Radden, Günter. 1985. Spatial metaphors underlying prepositions of causality. In Wolf Paprotté & René Dirven (eds.), *The ubiquity of metaphor: Metaphor in language and thought* (Current Issues in Linguistic Theory 29), 177–207. Amsterdam: John Benjamins. DOI: 10.1075/cilt.29.10rad.

Rappaport Hovav, Malka & Beth Levin. 2001. An event structure account of English resultatives. *Language* 77(4). 766–797. DOI: 10.1353/lan.2001.0221.

Redslob, M. Gustavus Mauritius. 1835. *De particulae Hebraicae* כי *origine et indole* [On the origin and nature of the Hebrew particle כי]. Leipzig: Weinedel.

Ringgren, Helmer. 2001. עָמַד *ʿāmaḏ*. In G. Johannes Botterweck, Helmer Ringgren & Heinz-Josef Fabry (eds.), *Theological dictionary of the Old Testament*, vol. 11, 16 vols., 178–187. Grand Rapids, MI: William B. Eerdmans.

Rissanen, Matti. 1991. On the history of *that*/zero as object clause links in English. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 272–289. New York, NY: Routledge.

Ritter, Elizabeth & Martina Wiltschko. 2019. Nominal speech act structure: Evidence from the structural deficiency of impersonal pronouns. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 64(4). 709–729. DOI: 10.1017/cnj.2019.10.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: Handbook in generative syntax* (Kluwer International Handbooks of Linguistics 1), 281–337. Dordrecht: Springer. DOI: 10.1007/978-94-011-5420-8_7.

Roberts, Craige. 2012 [1996]. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5(6). 1–69. DOI: 10.3765/sp.5.6.

Roberts, Ian. 2019. *Parameter hierarchies and Universal Grammar* (Rethinking Comparative Syntax). Oxford: Oxford University Press. DOI: 10.1093/oso/9780198804635.001.0001.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist aproach to grammaticalization* (Cambridge Studies in Linguistics 100). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486326.

Rodriguez, Daniel L. 2017. מ/לפני, אחר, *and* תחת*: An embodied cognitive approach to the Biblical Hebrew prepositions*. University of Stellenbosch dissertation.

Rooryck, Johan. 2003. The morphosyntactic structure of articles and pronouns in Dutch. In Jan Koster & Henk van Riemsdijk (eds.), *Germania et alia: A linguistic webschrift for Hans den Besten*. Groningen: University of Groningen.

Rooryck, Johan. 2019. 'Recycling' evidentiality: A research program. In Metin Bağrıaçık, Anne Breitbarth & Karen De Clercq (eds.), *Mapping linguistic data: Essays in honour of Liliane Haegeman*, 242–261. Ghent.

Rosenbaum, Peter Steven. 1965. *The grammar of English predicate complement constructions*. Massachusetts Institute of Technology dissertation.

Ross, John Robert. 1970. On declarative sentences. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 222–277. Waltham, MA: Ginn and Company.

Schoors, Antoon. 1981. The particle כי. *Oudtestamentische Studiën* 21. 240–276. DOI: 10.1163/9789004494039_017.

Seppänen, Aimo, Rhonwen Bowen & Joe Trotta. 1994. On the so-called complex prepositions. *Studia Anglica Posnaniensia* 29. 3–29.

Sloman, Steven A., Philip M. Fernbach & Scott Ewing. 2012. A causal model of intentionality judgment. *Mind & Language* 27(2). 154–180. DOI: 10.1111/j.1468-0017.2012.01439.x.

Sollamo, Raija. 2003. The passive with an agent in Biblical Hebrew and its rendering in the Septuagint. In Martin F. J. Baasten & Willem T. Van Peursen (eds.), *Hamlet on a hill: Semitic and Greek studies presented to professor T. Muraoka on the occasion of his sixty-fifth birthday* (Orientalia Lovaniensia Analecta 118), 617–629. Leuven: Peeters.

Sornicola, Rosanna. 1996. Alcune strutture con pronome espletivo nei dialetti Italiani meridionali [Some structures with expletive pronouns in Southern Italian dialects]. In Paola Benincà, Guglielmo Cinque, Tullio de Mauro & Nigel Vincent (eds.), *Italiano e dialetti nel tempo: Saggi di grammatica per Giolio C. Lepschy* [Italian and dialects over time: Essays in grammar for Giolio C. Lepschy], 323–340. Rome: Bulzoni Editore.

Spang-Hanssen, Ebbe. 1963. *Les prépositions incolores du français moderne* [The colorless prepositions of Modern French]. Copenhagen: University of Copenhagen dissertation.

Speas, Peggy & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria Di Sciullo (ed.), *Syntax and semantics*. Vol. 1: *Asymmetry in grammar* (Linguistics Today 57), 315–344. Amsterdam: John Benjamins. DOI: 10.1075/la.57.15spe.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Pragmatics* (Syntax and Semantics 9), 315–332. New York, NY: Academic Press. DOI: 10.1163/9789004368873_013.

Stalnaker, Robert. 2002. Common Ground. *Linguistics and Philosophy* 25(5/6). 701–721. JSTOR: 25001871.

Staps, Camil. 2018. A case study of reciprocal middles in Biblical Hebrew: The niphal of לחם. *Orientalia* 87(2). 159–183.

Staps, Camil. 2022. *Lipnē* 'in the face of': A Locative preposition with a threatening connotation. *Journal of Hebrew Scriptures* 22(1). 1–24. DOI: 10.5508/jhs29582.

Staps, Camil. 2023a. *Uses of Biblical Hebrew* בְּ bə *and* מִן min *in Genesis, Exodus, Joshua, Samel, Ruth, Ecclesiastes, and Esther* [data set]. DOI: 10.5281/zenodo.8314905.

Staps, Camil. 2023b. *Uses of Biblical Hebrew* כִּי kī *in Genesis, Judges, Samuel, and Ruth* [data set]. DOI: 10.5281/zenodo.8314818.

Staps, Camil & Martijn Beukenhorst. 2024. Semantic properties of prepositions: The distinction between causal *min* 'from' and *b*^{e} 'in'. *Brill's Journal of Afroasiatic Languages and Linguistics* 16(2). DOI: 10.1163/18776930-01602002.

Staps, Camil & Johan Rooryck. 2023. The interpretation of [+distal] in demonstratives and complementizers. *Linguistics* 61(5). 1195–1231. DOI: 10.1515/ling-2022-0178.

Staps, Camil & Johan Rooryck. 2024. Formalizing spatial-causal polysemy of Agent prepositions. *Semantics and Pragmatics* 17(4). DOI: 10.3765/sp.17.4.

Stirling, Lesley & Rodney Huddleston. 2002. Deixis and anaphora. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 1449–1564. Cambridge: Cambridge University Press. DOI: 10.1017/9781316423530.018.

Storms, Godfrid. 1966. *That*-clauses in Modern English. *English Studies* 47. 249–270. DOI: 10.1080/00138386608597258.

Straub, Sylvia A. 1974. The passive prepositions *de* and *par*. *The French Review* 47(3). 583–593. JSTOR: 387100.

Strauss, Susan. 2002. *This*, *that*, and *it* in spoken American English: A demonstrative system of gradient focus. *Language Sciences* 24(2). 131–152. DOI: 10.1016/S0388-0001(01)00012-2.

Strecker, Ivo. 1993. Cultural variations in the concept of 'face'. *Multilingua* 12(2). 119–142. DOI: 10.1515/mult.1993.12.2.119.

Suñer, Margarita. 1981. *Por* vs. *de*: Agential prepositions? *Hispania* 64(2). 278–283. DOI: 10.2307/341864.

Talmy, Leonard. 1988. Force dynamics in language and cognition. *Cognitive Science* 12. 49–100. DOI: 10.1207/s15516709cog1201_2.

Testen, David. 1998. The derivational role of the Semitic N-stem. *Zeitschrift für Assyriologie* 88. 127–145. DOI: 10.1515/zava.1998.88.1.127.

Thompson, Sandra A. 2002. "Object complements" and conversation. Towards a realistic account. *Studies in Language* 26(1). 125–163. DOI: 10.1075/sl.26.1.05tho.

Thompson, Sandra A. & Anthony Mulac. 1991. The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics* 15. 237–251. DOI: 10.1016/0378-2166(91)90012-M.

Thorion, Yochanan. 1984. *Studien zur klassischen hebräischen Syntax* [Studies on Classical Hebrew syntax] (Marburger Studien zur Afrika- und Asienkunde: Serie B, Asien 6). Berlin: Dietrich Reimer.

Trotzke, Andreas & Xavier Villalba. 2021. Expressive insubordination: A cross-linguistic study on *that*-exclamatives. In Andreas Trotzke & Xavier Villalba (eds.), *Expressive meaning across linguistic levels and frameworks*, 108–120. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198871217.003.0006.

Tsoulas, George. 1996. The nature of the subjunctive and the formal grammar of obviation. In Karen T. Zagona (ed.), *Grammatical theory and Romance languages: Selected papers from the 25th Linguistic Symposium on Romance Languages (LSRL XXV), Seattle, 2–4 March 1995* (Current Issues in Linguistic Theory 133), 293–306. Amsterdam: John Benjamins. DOI: 10.1075/cilt.133.23tso.

Tsumura, David T. 2014. Textual corruptions, or linguistic phenomena? The cases in 2 Samuel (MT). *Vetus Testamentum* 64. 135–145. DOI: 10.1163/15685330-12301140.

Tyler, Andrea & Vyvyan Evans. 2003. *The semantics of English prepositions: Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486517.

Ukosakul, Margaret. 2005. The significance of 'face' and politeness in social interaction as revealed through Thai 'face' idioms. In Robin T. Lakoff & Sachiko Ide (eds.), *Broadening the horizon of linguistic politeness* (Pragmatics & Beyond New Series 139), 117–125. Amsterdam: John Benjamins. DOI: 10.1075/pbns.139.11uko.

Van der Merwe, Christo H. J. 1993. Old Hebrew particles and the interpretation of Old Testament texts. *Journal for the Study of the Old Testament* 18(60). 27–44. DOI: 10.1177/030908929301806002.

Van der Merwe, Christo H. J., Jacobus A. Naudé & Jan H. Kroeze. 2017. *A Biblical Hebrew reference grammar*. 2nd edn. London: T&T Clark.

Van Wolde, Ellen. 2009. *Reframing biblical studies: When language and text meet culture, cognition, and context*. Winona Lake, IN: Eisenbrauns.

Van Wolde, Ellen. 2019. The niphal as middle voice and its consequence for meaning. *Journal for the Study of the Old Testament* 43(3). 453–478. DOI: 10.1177/0309089217743160.

Van Wolde, Ellen. 2021. *Nifʿal* verbs in the book of Genesis and their contribution to meaning. In *New perspectives in Biblical and Rabbinic Hebrew* (Cambridge Semitic Languages and Cultures 7), 431–454. Open Book Publishers. DOI: 10.11647/OBP.0250.15.

Vardi, Ruti. 2015. Favor: A construction of affection in Biblical Hebrew. *Hebrew Studies* 56(1). 49–69. DOI: 10.1353/hbr.2015.0025.

Visan, Florentina. 2000. The nature and status of exclamatives in Chinese. University of Bucharest unpublished manuscript.

Vriezen, Theodorus C. 1958. Einige Notizen zur Übersetzung des Bindewortes *Kī* [Some notes on the translation of the connective *kī*]. In Johannes Hempel & Leonhard Rost (eds.), *Von Ugarit nach Qumran* [From Ugarit to Qumran] (Beihefte zur Zeitschrift für die alttestamentliche Wissenschaft 77), 266–273. Berlin: De Gruyter. DOI: 10.1515/9783112314982-022.

Waltke, Bruce K. & Michael Patrick O'Connor. 1990. *An introduction to Biblical Hebrew syntax*. Winona Lake, IN: Eisenbrauns.

Walzer, Michael. 1992. The idea of holy war in ancient Israel. *The Journal of Religious Ethics* 20(2). 215–228. JSTOR: 40015152.

Watts, J. Wash. 1964. *A survey of syntax in the Hebrew Old Testament*. Grand Rapids, MI: William B. Eerdmans.

Weinert, Regina. 2012. Complement clauses in spoken German and English: Syntax, deixis and discourse-pragmatics. *Folia Linguistica* 46(1). 233–265. DOI: 10.1515/flin.2012.8.

Wierzbicka, Anna. 1988. *The semantics of grammar* (Studies in Language Companion Series 18). Amsterdam: John Benjamins. DOI: 10.1075/slcs.18.

Williams, Alexander. 2015. *Arguments in syntax and semantics* (Key Topics in Syntax). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139042864.

Williams, Ronald J. 2007. *Williams' Hebrew syntax*. John C. Beckman (ed.). 3rd edn. Toronto: Universiy of Toronto Press.

Yaguchi, Michiko. 2001. The function of the non-deictic *that* in English. *Journal of Pragmatics* 33(7). 1125–1155. DOI: 10.1016/S0378-2166(00)00045-X.

Yap, Foong Ha, Karen Grunow-Hårsta & Janick Wrona. 2011. Introduction: Nominalization strategies in Asian languages. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages* (Typological Studies in Language 96), 1–57. Amsterdam: John Benjamins. DOI: 10.1075/tsl.96.01yap.

Yu, Ning. 2001. What does our face mean to us? *Pragmatics & Cognition* 9(1). 1–36. DOI: 10.1075/pc.9.1.02yu.

Zanuttini, Raffaella & Paul Portner. 2003. Exclamative clauses: At the syntax-semantics interface. *Language* 79(1). 39–81. DOI: 10.1353/lan.2003.0105.

Zevakhina, Natalia. 2013. Syntactic strategies of exclamatives. *Eesti ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 4(2). 157–178. DOI: 10.12697/jeful.2013.4.2.09.

Zumthor, Paul & Walther von Wartburg. 1947. *Précis de syntaxe du français contemporain* [Syntax of contemporary French]. Bern: Francke.

Zwarts, Joost & Yoad Winter. 2000. Vector space semantics: A model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information* 9(2). 169–211. DOI: 10.1023/A:1008384416604.

Samenvatting

In deze dissertatie betoog ik op basis van taalkundige data dat mensen abstracte relaties vaak op een ruimtelijke manier conceptualiseren. De argumentatie hiervoor is gebaseerd op het gebruik van de functionele lexemen die zulke relaties uitdrukken, zoals het voorzetsel *uit*:

(1) a. *Opgeschrikt door de gasten vloog de kat de kamer **uit**.*
b. *De kat vloog weg **uit** angst voor de gasten.*

Mijn stelling is dat het niet toevallig is dat hetzelfde voorzetsel gebruikt wordt voor het uitdrukken van een ruimtelijke relatie enerzijds (1a) en een causale relatie anderzijds (1b). In deze dissertatie presenteer ik manieren om de polysemie van functionele lexemen als *uit* formeel te beschrijven.

Deel I behandelt causale relaties. In hoofdstuk 1 bestudeer ik met Johan Rooryck de Franse voorzetsels *de* '(weg) van' en *par* 'doorheen, via'. Deze voorzetsels hebben naast een ruimtelijke ook een causale betekenis:

(2) *Elle est adorée **de**/**par** son grand-père.*
'Ze is geliefd bij/wordt geliefd door haar grootvader.'

Er is een betekenisverschil tussen het gebruik van *de* en *par* dat is te verklaren op basis van het verschil in *afstand* dat deze voorzetsels uitdrukken: ruimtelijk gesproken geeft *de* een (verder afliggende) oorsprong aan, terwijl *par* de route beschrijft die wordt afgelegd. Voor het causale gebruik betekent dit dat het object van *de* minder agentief is in de situatie. In (2) zou de implicatie van *par* bijvoorbeeld kunnen zijn dat de grootvader zijn kleindochter verwent, terwijl het bij *de* om een mentale toestand gaat zonder gevolgen. Aangezien dit verschil is te herleiden op een verschil in ruimtelijke betekenis stellen we voor dat deze voorzetsels een algemene, abstracte betekenis hebben met een ruimtelijke component. De verschillende concrete betekenissen worden direct van deze abstracte betekenis afgeleid, afhankelijk van de manier waarop het domein (bv. dat van causaliteit) ruimtelijk

wordt geconceptualiseerd. We presenteren een semantisch formalisme om de afleiding van concrete betekenissen te beschrijven, gebruikmakend van een polymorf type $\langle \eta, \langle \theta, t \rangle \rangle$. Hierin worden η en θ door concrete types geïnstantieerd afhankelijk van de syntactisch-semantische context. Dit leidt tot een economischere analyse van het gebruik van zulke voorzetsels in passieve zinnen als (2), omdat de betekenis van deze voorzetsels in dit formalisme kan worden gerelateerd aan hun ruimtelijke betekenis.

Hoofdstuk 2, waarvan Martijn Beukenhorst co-auteur is, beschrijft een soortgelijk contrast tussen de Bijbels Hebreeuwse voorzetsels מִן *min* 'weg van, uit' en בְּ *bə* 'in, bij, met'. Ook hier is de causale betekenis af te leiden van de ruimtelijke betekenis. Het voorzetsel *bə* beschrijft een locatie in of dicht bij het object, wat er bijvoorbeeld voor zorgt dat *bə* gebruikt kan worden om instrumenten te markeren ('Als een man zijn slaaf of slavin *met* (*bə*) een staf slaat …'; Exod. 21:20). Een instrument is fysiek betrokken bij een handeling en staat er in die zin "dichtbij". *Min* daarentegen beschrijft een oorsprong en wordt daarom eerder gebruikt om degene die een beslissing heeft gemaakt aan te duiden ('het was niet [de wil] *van* (*min*) de koning'; 2 Sam. 3:37). Hierbij wordt de afstand uitgedrukt door *min* geïnterpreteerd als *dominantie* over de situatie: een agens die door *min* wordt geïntroduceerd heeft meer controle over de situatie dan een agens gemarkeerd met *bə*.

Deel II behandelt relaties tussen sprekers aan de ene kant en hetgeen besproken wordt aan de andere kant. Ook hier speelt *afstand* een belangrijke rol. In hoofdstuk 3 presenteren Johan Rooryck en ik een uniforme analyse van aanwijzende voornaamwoorden (*this*/*that* in [3a]) en onderschikkende voegwoorden gebaseerd op [+distal] voornaamwoorden (*that* in [3b]).

(3) a. *I have read **this**/**that** book.*
b. *I know **that** it is a good book.*

We laten zien dat aan gevallen als (3b) een ruimtelijke representatie van de *discourse space* ten grondslag ligt. Het voegwoord *that* plaatst de inhoud van de ingebedde zin "dicht bij" de hoorder (en dus "ver" van de spreker). Het gevolg is dat de inhoud gedeeld wordt met de hoorder en door haar moet worden aangenomen; de inhoud staat in de *Common Ground*. Zo kan (4b) alleen gebruikt worden als de hoorder weet dat bio-industrie nog is toegestaan, terwijl (4a) gebruikt kan worden om haar daarvan te overtuigen:

(4) a. *Bio industry is still allowed.*
b. ***That** bio industry is still allowed!*

In hoofdstuk 4 pas ik hetzelfde idee toe op het Bijbels Hebreeuwse voegwoord כִּי *kī*. Aan *kī* wordt traditioneel een groot aantal verschillende functies toegekend: het leidt lijdendvoorwerpszinnen in maar kan ook gebruikt worden als voegwoord van causaliteit ('omdat') of gevolg/doel ('(zo/op)dat'), tegenstellend voegwoord ('maar'), voorwaardelijk voegwoord ('gegeven dat, als, wanneer') en beperkend voegwoord ('behalve dat'). In dit hoofdstuk laat ik zien dat bijna al deze functies kunnen worden afgeleid van een algemenere functie, namelijk het uitdrukken van *Common Ground*. Deze functie is binnen het kader van hoofdstuk 3 uit te leggen wanneer men de algemeen geaccepteerde etymologie van *kī* volgt en aanneemt dat dit voegwoord teruggaat op een oorspronkelijk [+distal] deictisch morfeem.

Hoofdstuk 5 (deel III) ten slotte beschrijft het Bijbels Hebreeuwse voorzetsel לִפְנֵי *lipnē*, letterlijk 'voor het aangezicht van'. Volgens sommige auteurs kan dit voorzetsel in passieve zinnen met 'door' worden vertaald, zoals in 'Israël werd verslagen *door* (*lipnē*) de Filistijnen' (1 Sam. 4:2). Ik geef een aantal taalkundige argumenten tegen deze analyse, en stel in plaats daarvan voor dat *lipnē* een abstract-ruimtelijke betekenis heeft waarbij het 'aangezicht' staat voor iemands autoriteit en waardigheid. In de zinnen met נגף *ngp* 'verslaan' duidt *lipnē* de partij aan die autoriteit heeft over het onderwerp, maar dit wil niet zeggen dat deze partij niet direct betrokken is bij het fysieke 'verslaan'. 1 Sam. 4:2 kan beter worden vertaald met: 'Israël was uit het veld geslagen door de autoriteit van de Filistijnen'. Op basis van voorbeelden van dit werkwoord in actieve zinnen beargumenteer ik dat God de agens is van dit werkwoord en dat het dus God is die de uitkomst van veldslagen bepaalt: God bepaalt dat de Filistijnen autoriteit krijgen over Israël.

Samen leveren deze studies bewijs voor een diepgewortelde ruimtelijke conceptualisatie van abstracte relaties, in het bijzonder relaties van oorzaak en gevolg, relaties tussen sprekers en het besprokene, en relaties tussen mensen. De verschillende hoofdstukken laten bovendien zien dat binnen deze ruimtelijke conceptualisatie de notie *afstand* een belangrijke rol speelt.

Curriculum vitae

Camil Staps was born in 1996 in Leusden, the Netherlands. In 2019, he obtained a master degree in computing science *summa cum laude* from the Radboud University Nijmegen, specializing in foundations and implementations of the lambda calculus. An interest in the texts of the Hebrew Bible brought him to study Biblical Hebrew and exegesis under Ellen van Wolde at the same university. Sensing that a firmer linguistic grounding would be necessary to properly appreciate these texts, he decided to study Hebrew and Aramaic at Leiden University with Holger Gzella, finishing his research master *cum laude* in 2019. In the same year, he was awarded a *PhDs in the Humanities* grant by the Dutch Research Council to support his doctoral research into semantics and its interfaces with morphology and syntax in Biblical Hebrew. In this project, supervised by Johan Rooryck and Ellen van Wolde, he combined a traditional philological approach to Biblical Hebrew with formal semantic models based on testable data from modern languages. This dissertation presents the main results of that research.