

On the demonstrative nature of finite complementizers

Camil Staps and Johan Rooryck

August 8, 2022

Abstract

It is well-known that proximal *this* can be used to mark direct speech (*Sue said (this/*that): “It is raining”*) while distal *that* is required to mark indirect speech (*Sue said (*this/that) it is raining*). To our knowledge, there is no satisfactory account of this observation. We propose that the proximal/distal feature is recycled in the sentential domain to mark both *actual distance* (how similar the speech report is to the original utterance) and *Addressee involvement* (to what extent, and how, the Addressee is involved with the speech report). Addressee involvement in particular will be shown to account for a wider range of phenomena found in the distribution of overt and covert complementizers in a variety of European languages. This indicates that such complementizers are not neutral but carry an interpretive value. We then show that these results can be viewed as a generalisation of Peeters et al.’s (2021) theory of demonstrative reference. In this way, we can account for interpretive recycling of the proximal/distal distinction in both the nominal and the clausal domain.

Keywords: proximal/distal; presupposition; evidentiality; optional *that*; exclamative *that*; main clause *que/că*

1 Introduction

It is easy to see that the markers for direct and indirect speech are in complementary distribution:

- (1) a. *Sue said (this/*that): “It is raining.”*
b. *Sue said (*this/that) it is raining.* (Rooryck 2019: 257)

Admittedly, the syntactic status of *this* in (1a) is not the same as that of *that* in (1b). In (1a), *this* is a cataphoric pronoun coreferential with the direct discourse, while *that* in (1b) is commonly analysed as a complementizer introducing the embedded sentence. Still, the fact that a proximal and distal element are used for direct and indirect speech, respectively, requires an explanation. In this paper, we present a new, unified theory, with the goal to predict a broad range of ways in which the proximal/distal distinction is recycled in both demonstrative and complementation environments.

The standard view of the proximal/distal distinction in demonstratives is that it reflects physical distance. However, this view is untenable in the face of much experimental and corpus linguistic work, as Peeters et al. (2021) show in their review. We build on their framework for the study of demonstrative reference and aim to apply it to a broader range of uses of *this* and *that*.

The standard view on complementizers like *that* in (1b) is that they fulfil a primarily syntactic function and are largely void in terms of semantics and pragmatics, apart from carrying a feature indicating that they introduce a tensed rather than an untensed complement clause. In other words, mainstream theories of complementation do not ascribe meaning to the original distal semantics of the complementizer *that*. However, a number of studies have indicated that these complementizers carry additional interpretive information, mostly in 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 (1) also allows 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 adverb-COMP constructions in Romance (to be exemplified below). We then show how these recurring properties can be explained with reference to a proximal/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*.¹ These terms are introduced and exemplified in the context of direct and indirect speech (section 2). We then move to complementizers more generally in section 3, showing how Addressee involvement in particular can explain alternations between overt and covert complementizers. Sections 4 and 5 are dedicated to show that the proximal/distal distinction is used in a similar way in demonstratives, namely that distal demonstratives tend to be used more when the Addressee is more involved in the conversation. Finally, in section 6 we draw parallels between different uses of complementizers on the one hand and different uses of demonstratives on the other, with the goal to show that the main difference between the two is syntactic, not functional.

2 Direct and indirect speech

As mentioned in the introduction, 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:²

¹ 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 reappropriated to mark features of a different category (here, actual distance and Addressee involvement). However, contrary to what may happen in case of grammaticalization, the original deictic meaning is not lost. This perspective 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). Rather, it yields a more economical, polysemous view of demonstratives as exercising chameleon-like, distinct but strongly related functions, that vary according to the syntactic context they are used in. Grammaticalization may be a useful framework to describe the original bleaching of purely physical deixis to general deixis in the very earliest phases of the life cycle of a demonstrative, but the reappropriation of general deixis for actual distance and Addressee involvement is more economically described as recycling.

² The use of *this* for direct speech is rather marked in English, but this pattern is found much less markedly in other languages. For instance, Classical Syriac by default introduces direct speech with the complementizer *d-* which is ultimately derived from a proximal demonstrative.

- (1) a. Sue said (*this/*that*): “*It is raining.*”
b. Sue said (**this/that*) *it is raining.* (Rooryck 2019: 257)

We argue that this pattern is not arbitrary, but is based on the recycling of the category of physical distance (proximal/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 involvement*. Both provide a link between physical distance and the direct/indirect speech distinction.

First, observe that direct and indirect speech reports differ in the degree to which the report is similar to the original utterance. That is, for a speaker to faithfully utter (1a), Sue’s utterance must have been very similar to *It is raining*. This is not the case with indirect speech: with the right context, (1b) may be uttered after Sue has said something like *Why is it always raining when I want to go out?* Since direct speech reports are more similar to the original utterance, they 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, phonological information needed to represent accents, accompanying gestures, propositional content, entailments, 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. This is analogous to defining a point in the physical world with x, y, and z coordinates and measuring the distance between two points.³ 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 descriptive 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).⁴ We will use the term ‘actual distance’ to refer to both the Euclidean distance in the physical world and in the multidimensional conceptual space where similarity of speech reports are 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. This is an interpretation of the ‘distance’ between the referent (Sue’s utterance) and the Speaker (of (1ab)).

³ This is similar to Paul Churchland’s notion of ‘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 colour 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. Colours can then be compared as similar or dissimilar by measuring the distance between them. Churchland proposes state spaces for different sensory systems, and suggests that concepts can be represented in a state space for language use and propositional knowledge, too (1986: 299–306), which is what we attempt to do here.

⁴ Throughout, we use the term ‘referent’ for the thing to which the deictic expression refers (cf. Maes et al. 2022). Note that 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).

A direct speech report as in (1a) is ‘close’ to the Speaker, because its use suggests that the Speaker, unlike the Addressee, has direct, reliable, knowledge of Sue’s utterance. By uttering (1a), 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. On the other hand, an indirect speech report as in (1b) does not imply that the Speaker has direct evidence for the utterance. Speaker and Addressee can then share this indirect evidence: the content is in the Common Ground. 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. Closeness to the Addressee is represented as distance from the Speaker, hence a distal element is used. In other words, the Common Ground is distal for the Speaker. We see the Common Ground not as a region encompassing Speaker and Addressee, but redefine it as the intersection of the collections of information content tracked by the interlocutors (including propositions, utterances, questions, ...). The Common Ground 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) and the information content shared with the Addressee (distal). In this way, although the speech report is positioned either close to or far from the Speaker, this is actually interpreted as far from or close to the Addressee, 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. The relation between direct speech and direct evidentiality was already observed by Clark & Gerrig (1990: 793–795), and compared to indirect speech by Rooryck (2019: 256–257).⁵

This treatment of speech reports and demonstratives 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: 132). 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.

⁵ 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 Common Ground while *this* refers to a ‘Speaker-personal’ Ground, i.e., to knowledge held by the Speaker but not shared by the Addressee. Therefore, for reasons outlined above, direct speech reports are not placed in the Common Ground but remain personal to the Speaker, while indirect speech reports are shared with the Addressee.

3 Presupposition effects

Having shown how the proximal/distal distinction is recycled to mark actual distance (recycled as descriptive similarity) and Addressee involvement in the context of speech reports, we now turn to cases where overt complementizers contrast with covert ones. In these cases there is no difference in terms of actual distance, but the notion of Addressee involvement does generalise.⁶ Our position will be that overt complementizers, which are historically based on non-proximal elements, markedly involve the Addressee. In particular, we analyse the examples below using the notion of Common Ground.⁷ When information content is in the Common Ground, it is shared with the Addressee, and therefore ‘far’ from the Speaker; when information content is not in the Common Ground, it is instead ‘close’ to the Speaker. The proximal/distal distinction is thus recycled to indicate the absence/presence of content in the Common Ground.

3.1 Exclamatives

We first look at main clauses with overt complementizers, which in many languages can get an exclamative reading.⁸

- (2) a. *That bio industry is still allowed!*
b. *That he should have left without asking me!*
(Quirk et al. 1985: 841 via Zevakhina 2013: 167)
- (3) *Att du hann med tå-get!* [Swedish]
COMP you do.PST PREP train.DF
‘(It is surprising,) that you caught the train!’ (Delsing 2010: 17 via Zevakhina 2013: 167)
- (4) *Że też tyś potrafił coś takiego zrobić* [Polish]
COMP also you could something such do
‘That you could do something like this!’ (Storms 1966: 261)
- (5) *zaʕāqat̪ sədōm wa=ʕāmōrā kī rābb-ā wə=ħaʕtāt-ām kī*
kābəd-ā məʔōd [Biblical Hebrew]
outcry Sodom and=Gomorrah COMP great-3F.SG and=sin-3M.PL.POSS COMP
heavy-3F.SG very

⁶ We return to the reason why actual distance is not always relevant here in the conclusion.

⁷ We will use the term ‘Common Ground’ somewhat loosely here. Strictly speaking, the Common Ground is a set of propositions (Lewis 1969; Stalnaker 2002: 706). For us, Common Ground also needs to include the information content of interrogatives (e.g., the implicit question for (2b)) and utterances (for (4)). Such types could be forced into the Common Ground by coercing them into propositions of the type “A asked if ...” or “B said: ‘...’”, but this seems unwieldy. It would be possible to store this information on a conversational scoreboard (Lewis 1979), which can be seen as an extension of the Common Ground. However, this notion is not very well-defined, so we will continue to use the term ‘Common Ground’ here, with the note that it should be understood to contain different types of information content (propositions and question) as well as utterances.

⁸ 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).

‘The outcry of (/against) Sodom and Gomorrah, how great it is! And their sin, how very grievous!’
 (Genesis 18:20)

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, however. 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 (2a) or *What a cute kitten!* in that it can be used to attempt to convince the Addressee of its propositional content. By contrast, the sentence in (2a) does not make an attempt at informing or convincing the Addressee of its propositional content, but actually presupposes it to be shared information in the Common Ground. The use of the complementizer is therefore crucial for the interpretation as an exclamative.

We adopt Zanuttini & Portner’s (2003) analysis of exclamatives as factive.⁹ On this view, exclamatives make reference to a proposition and relate a certain Speaker stance (surprise, anger, etc.) to it. The proposition itself is presupposed: the sentence in (2a), for example, cannot be used to convince an Addressee of the fact that bio industry is still allowed. We propose that the complementizer in these exclamatives anaphorically refers to the presupposed proposition.¹⁰

In these cases, the referent (the presupposition in the Common Ground) is always ‘far’ from the Speaker: we do not find exclamatives with a complementizer or other grammatical marker that is specified for proximity.¹¹ In our framework, it is easy to see why: if, according to Zanuttini & Portner (2003), exclamatives require presupposition, they must refer to the Common Ground. An exclamative cannot at the same moment introduce new, Speaker-personal information content into the discourse. As a result, the information content must be close to the Addressee, and therefore a distal element must be used.

3.2 The *that*/null alternation in English object clauses

Finally, we can use Addressee involvement to explain the alternation between overt and covert complementizers in English object clauses. Consider (6):

⁹ 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 ...). The exact derivation of exclamatives is not important here; what is important is primarily the fact that exclamatives *are* factive.

¹⁰ Note that other syntactic strategies of exclamatives studied by Zevakhina (2013) also contain other anaphoric elements:

- | | | |
|------|--------------------------------------------------------------------------|-------------------------------------------------------------------|
| i. | <i>Zhège háizi!</i>
this child
‘What a child!’ | [Mandarin Chinese]

(Visan 2000: 9 via Zevakhina 2013: 169) |
| ii. | <i>It is so hot!</i> | (Michaelis 2001: 1040 via Zevakhina 2013: 166) |
| iii. | <i>Miša takoj bol’šoj!</i>
Miša such.NOM big.NOM
‘Misa is so big!’ | [Russian]

(Zevakhina 2013: 166) |

Although these anaphoric elements do not refer to the presupposed proposition, they still establish Common Ground between Speaker and Addressee.

¹¹ See the previous footnote, example (i) for a case where a proximal element can be used in an exclamative. But note that this is an exophoric/recognitional demonstrative and does not head the exclamative clause.

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

A common view is that the complementizer *that* in (6b) 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 null complementizer. Two in particular suggest that we are actually dealing with an interpretively meaningful alternation between an overt and a covert complementizer (as in the subsections above) and not with an entirely optional functional element.¹²

Firstly, Bolinger (1972: 58) already noticed that the sentence with *that* in (6b) suggests some context between Speaker and Addressee. This context may be extralinguistic, as in the scenario he sketches:

“Suppose you observe a stranger struggling to mount a tire. Feeling charitable you go over to him and say [6a]. Under these circumstances, [6b] would be inappropriate. But if the other person looks at you as if wondering why you came over, you might explain by saying [6b].” (Bolinger 1972: 58, example numbers adapted to ours)

In the words of Bolinger, the complementizer still “reflects the demonstrative character of *that*” (Bolinger 1972: 56) 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 (6b) (i.e., we assume there to be an implicit utterance along the lines of *Could you help me?*). The situation is then quite similar to that of exclamatives and the Romance adverb-COMP construction: the use of an overt complementizer signals content in the Common Ground. Again, then, the referent (the presupposed utterance) is analysed as ‘far’ from the Speaker, triggering a distal element, because the referent is in the Common Ground, close to the Addressee. In (6a), 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 Common Ground. The presence or absence of the complementizer thus marks the presence or absence of shared context in the Common Ground.

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 friendly, less emotive” than their counterparts with null complementizers

¹² 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. Kajzer-Wietrzny (2018) provides further evidence that structural complexity and the use of *that* over zero are positively correlated: she shows that *that* is used more frequently by interpreters and non-native speakers than by native speakers and suggests that this may be due to the cognitive load experienced by these groups; clarifying the sentence structure by using *that* may then help to reduce this cognitive load. 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 are style and register (Elsness 1984; Rissanen 1991), as well as Thompson & Mulac’s (1991) suggestion 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 likely.

(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” (1966: 263). Later, when it is important that objective facts are established, questions with *that* are used (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 the Common Ground. 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.¹³

3.3 The adverb-COMP construction in Romance

Presupposition and the Common Ground also play a role in the following type of construction, found in several Romance languages. This construction is formed by a sentence-initial adverb followed by an overt complementizer:¹⁴

- (7) *Evidentemente (que) Julia está muy enfadada* [Spanish]
 obviously COMP Julia is very angry
 ‘Obviously, Julia is very angry.’ (Etxepare 1997: 98–99 via Hernanz 2007: 165–166)

As is the case with exclamation marks, this complementizer engages in an interpretively meaningful alternation with a covert complementizer. According to Etxepare (1997: 98–99), the felicity of Spanish *que* in sentences like (7) is conditioned by the occurrence of the complement clause in preceding discourse. Thus, (7) is felicitous after another Speaker has uttered (8a), but not after they have uttered (8b).

- (8) a. *Creo que Julia está muy enfadada* [Spanish]
 b. *Creo que Julia está muy enojada*
 believe.1SG COMP Julia is very angry
 ‘I believe Julia is very angry’

It seems that this restriction of lexical identity is relatively strict, but not absolute. For instance, Pérez & Verdecchia (2022) describe cases of clausal doubling which likewise require lexical identity in the verb, even to the point that replacement by apparent synonyms like *enojarse* for *enfadarse* ‘to get angry’ is rejected (8a), while replacement of NPs by pronouns is nevertheless accepted, as in (9b):

- (9) a. **Que Juan se enoja, se enfadó* [Spanish]
 COMP Juan REFL got_mad.3SG REFL got_mad.3SG
 ‘As for Juan getting mad, he *did* get mad’ (Pérez & Verdecchia 2022)

¹³ 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’.

¹⁴ It may not be immediately obvious that the complementizers discussed here contain a non-proximal deictic element. We return to this issue below.

- b. *Que leyó el libro, seguro que lo leyó* [Spanish]
 that read.3SG the book, sure that it read.3SG
 ‘As for her reading the book, she read it for sure’ (Pérez & Verdecchia 2022)

Pérez & Verdecchia (2022) analyse this type of clausal doubling as follows: the first clause topicalizes a certain proposition as the Question Under Discussion (QUD); the second clause answers this QUD. This analysis transfers to (7) above: a linguistic antecedent like (8a) as the referent establishes the QUD to which (7) replies. It is clear, then, that *que* is licensed by the existence of a proposition in the Common Ground (namely, the polar QUD). This also explains why (7) is not felicitous after (8b). Setting aside the lexical identity requirement, the pragmatic conditions for the use of *que* are then very similar to those on *that* in exclamatives: *que* marks the existence of Common Ground and hence Addressee involvement.

In other Romance languages, the same construction receives an evidential interpretation.¹⁵

- (10) *Sigur (că) va veni* [Romanian]
 sure COMP will.3SG come
 ‘Of course s/he’s coming.’ (Cruschina & Remberger 2017: 89)

In (10), Romanian *că* may only be used when the Addressee could have inferred the propositional content of the complement. 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, if we suppose that both Ioana and Alexandru know that Ion is a big fan of the conference and would never miss it, 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 Common Ground, and hence Addressee involvement, is marked by an overt complementizer.

The situation in Neapolitan is again slightly different. The following contrast is discussed by Sornicola (1996: 334–336) and Ledgeway (2011: 286–289):

- (11) a. *Chello_i s’è astutato* [‘o riscaldamento]_j [Neapolitan]
 that.N self=is turned_off the.M.SG heating.M
 ‘(The fact is/Because) the heating has gone off.’
 b. *Chillo_i s’è astutato* [‘o riscaldamento]_i [Neapolitan]
 that.M self=is turned_off the.M.SG heating.M
 ‘The heating has gone off.’ (Ledgeway 2011: 286¹⁶)

On Ledgeway’s double subject analysis, *chello/chillo* is not a complementizer but a demonstrative. In (11b), it is coreferential with the second subject (“It has gone off, the heating”). In (11a) it is not, and

¹⁵ We use an example in Romanian here; for examples from other languages see Cruschina & Remberger (2017). For more discussion on Romanian, see Hill (2012).

¹⁶ The indices in (10b) have been corrected from the source after consultation with Adam Ledgeway (p.c., June 16, 2022).

obligatorily takes neuter gender. 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 (11a) the demonstrative refers to this associated presupposition, as is the case with the complementizers in Spanish and Romanian.

We should pause here for a moment to reflect on the origin of these complementizers. Above, we argued that the complementizer *that* marks the use of Common Ground because the Common Ground includes the Addressee and is therefore ‘far’ from the Speaker. However, the Romance complementizers here are not demonstrative synchronically, so how does a proximal/distal feature fit in? First, we should point out that these complementizers derive from Latin *quod*, which is composed of an interrogative element *qu-* and the originally neuter medial demonstrative *id*. Therefore, these complementizers do diachronically derive from a non-proximal demonstrative element. Furthermore, there is reason to believe that the interrogative element *qu-* is incompatible with proximity. For instance, note that English has *what* from *that* and *where* from *there*, but not *whis* from *this* or *where* with an /i/-vowel from *here*. Rooryck (2003: 11–12) suggests that this is because something that is proximate to the Speaker is necessarily known to them.

Our analysis of these complementizers is very similar to that of exclaimatives. In (10), 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 exclaimatives” and analyses them as in (12a):

- (12) a. [Force Adv/A_[+evident]] [Focus [+f]] [Topic COMP ...]]
 b. ¡Claro que te lo voy a dar! [Spanish]
 clear that to_you it go to give
 ‘Of course I will give it to you!’

(Gutiérrez-Rexach 2001: 184–185)

In (12b), the evidential adverb *claro* requires that its complement makes reference to some question under discussion (QUD). For example, (12b) 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. This example thus illustrates the division of labour between the sentence-initial adverb and the complementizer and the parallel with the exclaimatives discussed above. As with exclaimatives, the function of the complementizer is to mark the existence of Common Ground between Speaker and Addressee, while the sentence-initial

adverb only specifies the evidential interpretation.¹⁷ The distal element is therefore again used to signal Addressee involvement.

3.4 Presupposition effects: summary

To summarise the findings from this section: evidence from a variety of constructions (exclamatives, the adverb-COMP construction in Romance, and English ‘optional’ *that*) suggests that the alternation between an overt complementizer with a non-proximal feature and a covert complementizer is related to presupposition. We explain this by suggesting that the complementizer refers to a proposition in the Common Ground. Distal elements are used in these complementizers because the Common Ground includes the Addressee, who is ‘far’ from the Speaker. Note that the theory correctly predicts that we do not find proximal elements in these environments. These would correspond to presuppositions that are not shared with the Addressee; a contradiction in terms, since presuppositions are necessarily shared by all interlocutors. The only available alternation is with a covert complementizer, which marks the absence of a proposition in the Common Ground.

Our analysis raises questions for the traditional account of the grammaticalization of *that* and cognate complementizers. On this account, *that* became a complementizer as a result of reanalysis of a cataphoric demonstrative: *I say that: he comes* becomes *I say that he comes* (e.g. Diessel 1999: 123–125). The problem with this is that a cataphoric demonstrative does not refer to something in the Common Ground, but introduces new information, while our analysis suggests that the complementizer *that* does refer to Common Ground. 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 Bate in preparation for a generalisation to Indo-European). In such a construction, the first pronoun introduces new information but the second can be seen as referring to the Common Ground (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 we focus here 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: *Sážaljavam, deto ne možax da doйда* ‘I regret that I couldn’t come’ (Krapova 2010: 1240). The crucial factor thus seems to be not the syntactic category but the deictic feature of the element. A good starting point to look for further parallels in other languages would be Bate’s (in preparation) survey of the origin of finite complementizers in Indo-European.

¹⁷ Note that the term “Speaker-oriented” for these evidential sentences thus 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 yields the odd situation that a distal element, which is typically used to trigger a more objective interpretation by placing something in the Common Ground, actually generates a Speaker-oriented reading.

4 Exophoric demonstratives

In the previous two sections we examined the complementizer *that*. We compared this functional element to both the proximal cataphoric demonstrative *this* (for direct speech) and a covert complementizer (in main and object clauses). Both sections were concerned with reference to information content, namely, the meaning of utterances that may or may not be in the Common Ground. 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 5).

Demonstratives are exophoric when they refer to entities “in the speech situation” (Diessel 1999: 93).¹⁸ 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 be one of the physical distance between the referent and the deictic origo (typically, the Speaker). However, a wealth of experimental results have shown this view to be too simplistic (Peeters et al. 2021). Physical aspects of the relation between Speaker and referent (which are not limited to distance, but can also include things like visibility, whether the referent is uphill or downhill, etc.) are one factor, but they do not solely determine 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).¹⁹ 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 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 of *este* encompasses the entire conversational dyad, including both Speaker and Addressee. However, 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 maximise information density. In natural language, however, psychological factors are more important, which explains the different results found by Jungbluth (2003).

¹⁸ It should be noted that anaphoric demonstratives could also be said to refer to entities in the speech situation, but only indirectly, via an antecedent in the surrounding discourse.

¹⁹ The choice of a demonstrative also depends on referent-intrinsic factors like animacy and grammatical gender, but these are not relevant to us here.

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 the Common Ground, while *this* refers to a private zone around the Speaker. These psychological factors can be further interpreted pragmatically. Consider the following example:

(13) *How’s that throat?* (Lakoff 1974 via Cheshire 1996: 376)

The demonstrative in (13) 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* not only signals this previous knowledge but also expresses Speaker involvement which can be interpreted as empathy with the Addressee. 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.

By contrast, physical factors correspond to actual distance, i.e. the recycling of the spatial relation between the deictic expression and its referent. To see why, consider that exophoric demonstratives are often if not always accompanied by a pointing gesture, and can even be replaced by one (Jouitteau 2004: 109). We take this as an indication that the demonstrative has a position in the physical world, like the referent, so that the relationship between the two is determined by physical factors like distance or visibility.

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:

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

‘The lawyer talked to a client. Since he didn’t have much time, they agreed to have another meeting next week.’ (Diessel 1999: 96)

Unlike the personal pronoun *er*, the demonstrative pronoun *der* can only be coreferential with *einem Klienten* ‘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:²⁰

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

²⁰ This is part of what Diessel (1999: 100–105) calls the discourse deictic use of demonstratives. However, we only include references to propositions (*That’s false*), not references to illocutions (*That’s a lie*).

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 (16b) while their exophoric counterparts can (16a):

- (16) 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 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. (2021) recognize that the main difference between these types of corpora is the type of interaction between Speaker (writer) and Addressee. In spoken dialogue, there is continuous feedback from the Addressee to the Speaker. As a result, the Speaker can be relatively 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 a shared Common Ground 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 seems to prevent the writer from assuming that they share proper Common Ground with the Addressee. This means that scientific authors will more frequently assume that their readers do not share Common Ground with them, and hence use proximal demonstratives more frequently.²¹

²¹ This suggestion generates falsifiable hypotheses that can be tested against other types of corpora. For instance, we would expect unidirectional speeches 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 less clear preference than regular interactional discourse (Peeters et al. 2021: 421) can be explained this way, since evaluations are inherently personal and not in the Common Ground. 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.

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. We conclude with 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.

6 Conclusion

We have proposed a unified analysis of the recycling of the proximal/distal distinction between the demonstratives *this* and *that* in terms of *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 alternation between *this* and *that* to mark direct/indirect speech, and the alternation between *that* (or a parallel finite complementizer) and a covert complementizer in a variety of contexts. These abstract distances are interpreted in different ways depending on the type of referent, as shown in the following table:

	Entities (DP)	Information content (CP)
Exophoric	<p>Exophoric demonstratives (§4)</p> <p>Actual distance in the concrete physical world</p> <p>Addressee involvement: interpreted as psychological factors (psychological distance, joint attention, empathy, ...)</p>	<p>Direct / indirect speech (§2)</p> <p>Actual distance in a multidimensional conceptual world, interpreted as descriptive similarity</p> <p>Addressee involvement: interpreted as evidentiality; proximity is private witness evidentiality</p>
Anaphoric	<p>Anaphoric demonstratives (§5)</p> <p>Addressee involvement: <i>that</i> used over <i>this</i> to interact and empathise with the Addressee</p>	<p>Presupposition (§3)</p> <p>Addressee involvement: <i>that</i> used over \emptyset to signal content in the Common Ground</p>

The four types of environments discussed above have been categorised according to two binary properties here. First, our deictic elements refer to either information content or entities. We studied information content in sections 2 (direct and indirect speech) and 3 (presuppositions), and entities in sections 4 and 5 (exophoric and anaphoric demonstratives). Second, the well-known distinction between exophoric and anaphoric demonstratives for reference to entities generalises 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 presuppositions refer only indirectly (to entities via linguistic antecedents, and to information content through a mental model of the discourse state).

There are two gaps in this table. First, the ‘distance’ between referent and deictic expression does not seem to be 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. But a direct link between the referent and the deictic expression only exists in exophoric reference: in anaphoric reference, the link is indirect, through an intermediate linguistic entity. The fact that this link is intermediate 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.

Second, proximal elements appear to be incompatible with anaphoric reference to information content (presupposition). This gap has already been explained in section 3: using a proximal 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 Common Ground) without introducing it (because anaphoric reference is used). Such usage of language would be incompatible with cooperative conversation.

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:

(13) *How's that throat?*

(Lakoff 1974 via 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” (1966: 263). Thus, the demonstrative *that* in (13) would engage with the Addressee, whereas it is the absence of the complementizer *that* that does this for Storms (1966). By fleshing out what Addressee involvement really means in these different types of environments, the paradox can be resolved: Cheshire 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 generalisation from which other categories, 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. 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 cannot be precisely situated in space. The selection of a tensed complement by the complementizer *that*, which is often taken to be its only interpretive value, 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, we might wonder where the relativizer *that* fits in the table above. Its position is clearly in the lower left quadrant for anaphoric reference to entities. However, note that only the distal demonstrative can be used as a 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 Common Ground 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 Common Ground (as with reference to information content). We thus find the distinction between overt and covert relativizers to be similar to that between overt and covert complementizers. For example, (17a) 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 (17b), 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 rack), in contrast to the open-ended nature of (17a).²²

- (17) a. *There was nothing unusual \emptyset caught your eye when you came in?*
(Inspector Morse, season 7, episode 1)
b. *Was there anything that/? \emptyset caught your eye while browsing through the racks?*

In this paper we have analysed a number of high-frequency uses of the proximal/distal distinction, but our discussion has not been comprehensive. It is expected that other contexts will interpret actual distance and Addressee involvement in different ways. 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, the present paper presents the instrumentarium for further analysis of more specific kinds of reference.

References

- Axel-Tober, Katrin. 2017. 'The development of the declarative complementizer in German'. *Language* 93/2:e29–e65.
- Bate, Danny L. In preparation. 'Where did *that* come from? The distribution and development of neutral finite complementizers in Indo-European'.

²² Fox & Thompson (2007) give seven predictors for zero relativizers which can be subsumed under the notion of monoclausality. However, the examples in (17) hardly differ on their variables, suggesting that Addressee involvement may also be a relevant factor. Also cf. Bolinger's (1972: 70) discussion of *Is there anything (that) I can say?* A context where words seem inadequate (e.g. after a "terrible faux pas") demands for a zero relativizer. However, if you deliver a package for a friend and ask them if you should also deliver a message, *that* would be appropriate, since the event is clearly grounded and

- Beal, Joan. 1988. 'Goodbye to all "that"? The history and present behaviour of optional "that"'. In *An historic tongue: Studies in English linguistics in memory of Barbara Strang*, edited by John Honey and Graham Nixon, 49–66. London: Routledge.
- Biberauer, Theresa. 2017. 'Factors 2 and 3: A principled approach'. *Cambridge Occasional Papers in Linguistics* 10/3:38–65.
- Bolinger, Dwight. 1972. *That's That*. *Janua Linguarum* 155. The Hague: Mouton.
- Churchland, Paul M. 1986. 'Some reductive strategies in cognitive neurobiology'. *Mind* 95/379:279–309. Reprinted in Paul M. Churchland, 1992, *A neurocomputational perspective. The nature of mind and the structure of science*, 77–110. Cambridge, MA: The MIT Press.
- Clark, Herbert H., and Richard J. Gerrig. 1990. 'Quotations as Demonstrations'. *Language* 66:764–805.
- Çokal, Derya, Patrick Sturt, and Fernanda Ferreira. 2014. 'Deixis: *This* and *that* in written narrative discourse'. *Discourse Processes* 51:201–229.
- Coventry, Kenny R., Berenice Valdés, Alejandro Castillo, and Pedro Guijarro-Fuentes. 2008. 'Language within your reach: Near-far perceptual space and spatial demonstratives'. *Cognition* 108:889–895.
- Cruschina, Silvio, and Eva-Maria Remberger. 2017. 'Before the Complementizer: Adverb Types and Root Clause Modification'. In *Adjective Adverb Interfaces in Romance*, edited by Martin Hummel and Salvador Valera, 81–109. *Linguistics Today* 242. Amsterdam: John Benjamins.
- Delsing, Lars-Olof. 2010. 'Exclamatives in Scandinavian'. *Studia Linguistica* 64:16–36.
- Diessel, Holger. 1999. *Demonstratives. Form, function, and grammaticalization*. *Typological Studies in Language* 42. Amsterdam: John Benjamins.
- Ettxepare, Ricardo. 1997. 'The Grammatical Representation of Speech Events'. PhD diss., University of Maryland.
- Fox, Barbara A., and Sandra A. Thompson. 2007. 'Relative clauses in English conversation: Relativizers, frequency, and the notion of construction'. *Studies in Language* 31/2:293–326.
- Hernanz, M. Lluïsa. 2007. 'From Polarity to Modality. Some (a)symmetries between *bien* and *sí* in Spanish'. In *Coreference, Modality, and Focus: Studies on the Syntax-Semantics Interface*, edited by Luis Eguren and Olga Fernández Soriano, 133–169. *Linguistics Today* 111. Amsterdam: John Benjamins.
- Hill, Virginia. 2012. 'A main clause complementizer'. In *Main clause phenomena: New horizons*, edited by Lobke Aelbrecht, Liliane Haegeman, and Rachel Nye, 279–296. *Linguistics Today* 190. Amsterdam: John Benjamins.
- Jouitteau, Mélanie. 2004. 'Gestures as Expletives : Multichannel Syntax'. In *WCCFL 23 Proceedings*, edited by G. Garding and M. Tsujimura, 101–114. Somerville, MA: Cascadilla Press.
- Jungbluth, Konstanze. 2003. 'Deictics in the conversational dyad'. In *Deictic conceptualisation of space, time and person*, edited by Friedrich Lenz, 13–40. Amsterdam: John Benjamins.
- Kajzer-Wietrzny, Marta. 2018. 'Interpretese vs. Non-native Language Use: The Case of Optional *That*'. In *Making Way in Corpus-based Interpreting Studies*, edited by Mariachiara Russo, Claudio Bendazzoli, and Bart Defrancq, 97–113. *New Frontiers in Translation Studies*. Singapore: Springer.
- Krapova, Iliyana. 2010. 'Bulgarian relative and factive clauses with an invariant complementizer'. *Lingua* 120:1240–1272.
- Lakoff, Robin. 1974. 'Remarks on *this* and *that*'. In *Proceedings of the Tenth Regional Meeting of the Chicago Linguistics Society*, edited by Michael W. La Galy, Robert A. Fox, and Anthony Bruck, 321–344. Chicago: Chicago Linguistic Society.

- Ledgeway, Adam. 2011. 'Subject licensing in CP. The Neapolitan double-subject construction'. In *Mapping the left periphery: The cartography of syntactic structures*, edited by Paola Benincà and Nicola Munaro, 257–296. Oxford: Oxford University Press.
- Lewis, David. 1969. *Convention: A Philosophical Study*. Cambridge, Massachusetts: Harvard University Press.
- Lewis, David. 1979. 'Scorekeeping in a Language Game'. *Journal of Philosophical Logic* 8:339–359.
- Maes, Alfons, Emiel Kraemer, and David Peeters. 2022. 'Explaining variance in writer's use of demonstratives: A corpus study demonstrating the importance of discourse genre'. *Glossa: a journal of general linguistics* 7/1:1–36.
- Peeters, David, Emiel Kraemer, and Alfons Maes. 2021. 'A conceptual framework for the study of demonstrative reference'. *Psychonomic Bulletin & Review* 28:409–433.
- Pérez, Carlos Muñoz, and Matías Verdecchia. 2022. 'Information-based "island effects" in Spanish clausal doubling'. Presentation at the 45th Generative Linguistics in the Old World Colloquium (GLOW 45), April 27th.
- Quirk, Randolph, Greenbaum Sidney, Geoffrey Leech, and Jan Svartnik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rissanen, Matti. 1991. 'On the history of *that*/zero as object clause links in English'. In *English Corpus Linguistics. Studies in Honour of Jan Svartvik*, edited by Karin Aijmer and Bengt Altenberg, 272–289. New York: Routledge.
- Rooryck, Johan. 2003. 'The morphosyntactic structure of articles and pronouns in Dutch'. In *Germania et alia: A linguistic webschrift for Hans den Besten*. Groningen: University of Groningen. <http://www.let.rug.nl/koster/DenBesten/contents.htm>.
- Rooryck, Johan. 2019. "'Recycling" evidentiality: a research program'. In *Mapping linguistic data. Essays in honour of Liliane Haegeman*, edited by Metin Bağrıaçık, Anne Breitbarth and Karen De Clercq, 242–261. Ghent.
- Ross, John Robert. 1970. 'On declarative sentences'. In *Readings in English transformational grammar*, edited by Roderick A. Jacobs and Peter S. Rosenbaum, 222–277. Waltham, Massachusetts: Ginn and Company.
- Sornicola, Rosanna. 1996. 'Alcune strutture con pronome espletivo nei dialetti italiani meridionali'. In *Italiano e dialetti nel tempo. Saggi di grammatica per Giolio C. Lepschy*, edited by Paola Benincà, Guglielmo Cinque, Tullio de Mauro, and Nigel Vincent, 323–340. Roma: Bulzoni Editore.
- Stalnaker, Robert. 2002. 'Common Ground'. *Linguistics and Philosophy* 25:701–721.
- Stirling, Lesley, and Rodney Huddleston. 2002. 'Deixis and anaphora'. In *The Cambridge grammar of the English language*, edited by Rodney Huddleston and Geoffrey K. Pullum, 1449–1564. Cambridge: Cambridge University Press.
- Storms, G. 1966. 'That-clauses in Modern English'. *English Studies* 47:249–270.
- Strauss, Susan. 2002. 'This, that, and it in spoken American English: a demonstrative system of gradient focus'. *Language Sciences* 24:131–152.
- Thompson, Sandra A., and Anthony Mulac. 1991. 'The discourse conditions for the use of the complementizer *that* in conversational English'. *Journal of Pragmatics* 15:237–251.
- Trotzke, Andreas, and Xavier Villalba. 2021. 'Expressive insubordination: A cross-linguistic study on *that*-exclamatives'. In *Expressive Meaning Across Linguistic Levels and Frameworks*, edited by Andreas Trotzke and Xavier Villalba, 108–120. Oxford: Oxford University Press.
- Tsoulas, George. 1996. 'The nature of the subjunctive and the formal grammar of obviation.' In *Grammatical theory and Romance languages. Selected papers from the 25th linguistic symposium on Romance languages (LSRL XXV), Seattle, 2–4 March 1995*, edited by Karen Zagona, 293–306. Current issues in linguistic theory 133. Amsterdam: John Benjamins.

- Visan, Florentina. 2000. *The nature and status of exclamatives in Chinese*. Unpublished ms. University of Bucharest.
- Wierzbicka, Anna. 1988. *The semantics of grammar*. Studies in language companion series 18. Amsterdam: John Benjamins.
- Zanuttini, Raffaella, and Paul Portner. 2003. 'Exclamative Clauses: At the Syntax-Semantics Interface'. *Language* 79:39–81.
- Zevakhina, Natalia. 2013. 'Syntactic Strategies of Exclamatives'. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 4:157–178.