

Martina Wiltschko*

The grammar of self-talk. What different modes of talking reveal about language

<https://doi.org/10.1515/tl-2024-2024>

Abstract: Self-talk has played an important role in theorizing about the function of language in the psychological and philosophical literature. Linguistic investigations of self-talk, however, are scarce. It is shown that there are several modes of self-talk including (i) thinking out loud, which is characterized by the absence of an addressee and (ii) having a conversation with oneself, which is characterized by the presence of a grammatically represented addressee role. In the latter, the person engaged in self-talk may hold the role of the speaker or the addressee. Thus, the grammatical restrictions on self-talk serve as a hitherto underexplored window into the grammatical representation of speaker and addressee roles. Different models for the syntax at the top are compared and an argument is made for Wiltschko's Grammar of Interactional Language.

Keywords: self-talk; inner speech; performative hypothesis; speech act structure; interactional language

1 Introduction

The goal of this article is to explore the grammar of self-talk, where I take self-talk to be characterized by two defining properties, as in (1) (based on Latinjak et al. 2023).¹

- (1) The defining properties of self-talk
 - a. always consists of linguistic forms (i.e., is overtly realized)
 - b. the sender of the message is also the receiver

The first property distinguishes it from what is sometimes referred to as *inner speech*, which need not be formulated with specific linguistic means; the second property

¹ I use these defining properties in order to be precise about the object of study. It serves as a working definition, without theoretical status. As we shall see, there are various modes of self-talk which are more or less similar to inner speech on the one hand and typical conversations on the other hand.

***Corresponding author: Martina Wiltschko**, ICREA/Universitat Pompeu Fabra, Traducció y Ciencias del Lenguaje, Universitat Pompeu Fabra, Roc Boronat, 138, 08018 Barcelona, Spain,
E-mail: martina.wiltschko@upf.edu

distinguishes it from talk addressed to another person, which I refer to as a *typical conversation*. Note that I do not consider pretend conversations that happen in the absence of the intended addressee to be included in this definition of self-talk.

Self-talk has been extensively studied within psychology and philosophy, but research on the linguistics behind it is scarce, let alone its grammatical properties. This is even though language is at the core of the phenomenon. For example, in a recent overview article on self-talk, Latinjak et al. (2023: 355) start by emphasizing the critical role of language: “Human language is a unique phenomenon in nature that is used to communicate with other members of the species (Hockett 1959) and, to a similar extent, to communicate with oneself. This latter human behaviour is known as self-talk and has long fascinated researchers”. The extent to which specific linguistic properties of self-talk are addressed is limited. For example, this concerns the observation that people engaged in self-talk may refer to themselves with 1st or 2nd person pronouns. Strikingly, in one of the rare linguistic studies of self-talk, Holmberg (2010) demonstrates that the use of 1st and 2nd person pronouns in self-talk has linguistic implications, which are further explored in Ritter and Wiltschko (2021) and Goddard et al. (2022). However, these linguistic aspects do not inform the exploration of self-talk in the psychological literature, as evidenced by the fact that this work is not part of Latinjak et al.’s review, and neither is the only other linguistic contribution to the study of self-talk, namely that of Geurts (2018).

Thus, the linguistics of self-talk is an under-studied topic, even though it has the potential to serve as a window into several questions that pertain to the nature of human language. This is because self-talk presents us with a peculiar case that seems to situate the person engaging in self-talk somewhere in between thinking and communicating. It thus has the potential to contribute to the ancient question regarding whether the core function of language is thought or communication. In addition, this seemingly non-typical use of language allows us to compare recent proposals that syntacticize pragmatic aspects of language. Linguistic research over the past 20 years has seen an emerging consensus that properties of the speech act participants (i.e., speaker and addressee) are grammatically encoded (Giorgi 2010; Hill 2007; Zanuttini 2008). The empirical properties that can be modelled on this assumption relate to those, where aspects of language are sensitive to specific properties of the speaker or the addressee. For example, many languages display agreement with speech act participants, even if they do not serve as arguments within the proposition (Ross 1970). This is illustrated in (2) (Miyagawa 2012: 8) for Basque, where agreement on the final auxiliary is with the gender of the addressee.

- (2) a. Pettek lan egin dik
 Peter work do.PRFX AUX-2MASC
 ‘Peter worked.’

- b. Pettek lan egin din
 Peter work do.PRF AUX-2FEM
 'Peter worked.'

Assuming that agreement is syntactically conditioned, addressee agreement of this type provides striking evidence that speech act participants are indeed syntactically represented.² This conclusion is further corroborated by the fact that addressee agreement may be sensitive to clause-type and hence must be sensitive to syntactic properties (Miyagawa 2017; Oyharçabal 1993).

Several other phenomena support the hypothesis that speech act participants are part of the syntactic representation. For example, the nature of the relationship between speaker and addressee can affect aspects of grammar that are sensitive to formality, such as the T/V distinction in pronouns illustrated for German in (3).

- (3) a. **Frau Professor Zauner**, reden **Sie** mit sich selbst?
 'Ms. Professor Zauner, are you talking to yourself?'
 b. **Mathilde**, redest **du** mit dir selbst?
 'Mathilde, are you talking to yourself?'

The grammatical properties of this sensitivity can be modelled by encoding speaker and addressee features at the very top of the syntactic structure (Portner et al. 2019).

Similarly, there are linguistic means that are sensitive to the epistemic states of the speech act participants. For example, some German discourse particles are sensitive to the epistemic state of the addressee (or, more precisely, what the speaker assumes it to be). *doch* is used when the speaker assumes that the addressee knows (or should know) the propositional content of their utterance. In contrast, *nämlich* is used when the speaker assumes that the addressee does not (or cannot) know the propositional content. For example, when strangers meet each other on a plane, the speaker cannot assume that the addressee knows anything about their personal lives (such as whether they have a dog). Hence, in this context, *doch* is unacceptable while *nämlich* is acceptable, as shown in (4). In contrast, when the interlocutors know each other enough to know whether they have a dog, the reverse holds, as shown in (5). The data in (4) and (5) are from Upper Austrian German.

- (4) **Strangers** on a plane, at the start of the conversation
 I: San si vü untawegs?
 are you.FRML much around
 'Are you travelling a lot?'

2 There are approaches towards agreement according to which it is not necessarily a syntactically conditioned phenomenon (e.g. Wechsler and Zlatić 2003).

R1: *Na. I hob **doch** an Hund. Do kann I ned so leicht
No. I have DOCH a dog. there can I not so easily
weg.

away

‘No. (you know) I have a dog. So I cannot easily travel.’

R2: Na. I hob **nämlich** an Hund. Do kann I ned so
No. I have NÄMLICH a dog. there can I not so
leicht weg.

easily away

‘No. (you see) I have a dog. So I cannot easily travel.’

(5) **Acquaintances** where I knows R has a dog.

I: Wü-st du mit uns im summa noch Griechenland
want-2sg you with us in.the summer to Greece
kumma?

come

‘Do you want to come to Greece with us this summer?’

R1: Na. I hob **doch** an Hund. Do kann I ned so leicht
No. I have DOCH a dog. there can I not so easily
weg.

away

‘No. (as you know) I have a dog. So I cannot easily travel.’

R2: *Na. I hob **nämlich** an Hund. Do kann I ned so leicht
No. I have NÄMLICH a dog. there can I not so easily
weg.

away

‘No. (you see) I have a dog. So I cannot easily travel.’

One way to model the addressee orientation of discourse particles is to assume that they associate with a syntactically represented addressee role (Thoma 2016; Wiltschko 2024b).

These phenomena, among several others, provide empirical support for representing speech act participants in the layer(s) of syntactic structure that embed classic sentence structure, which I refer to as p(ropositional)-structure. This general idea is schematized in (6), abstracting away from details of individual analyses.

(6) [SA-participants [p-structure]]

In light of the structure in (6), the question arises as to what happens in self-talk. Is there still evidence for the syntactic presence of an addressee role when the speaker and the addressee are the same person? This is the core question I explore in this paper, and I argue that the answer depends on the mode of self-talk. Specifically, based on linguistic evidence, I argue, following Ritter and Wiltschko (2021) that there are (at least) two modes of self-talk, as summarized in (7).

- (7) Two modes of self-talk:³
- a. “*Thinking out loud*” involves a speaker/thinker only, but no addressee (role)
 - b. “*Having a conversation with oneself*” involves both a speaker and an addressee role where the person engaging in self-talk may assume either role

The linguistic properties of these modes of self-talk and the differences concerning typical conversations provide us with a novel window into the grammatical representation of the speech act participants. More broadly, it allows for a new way to approach two ancient questions: one regarding the relation between language, thought, and communication and the other regarding the phenomenon of self-talk itself. Neither of these questions has been addressed from the perspective taken here, i.e., the grammar of self-talk.

The paper is organized as follows. In Section 2, I introduce in more detail the phenomenon of self-talk and some of its linguistic properties. In Section 3, I introduce competing analyses for the syntax of the very top, where information about the speech act participants is encoded. In Section 4, I introduce an analysis of self-talk within the interactional spine hypothesis (Wiltschko 2021). In Section 5, I compare this analysis to alternatives within competing frameworks. Finally, in Section 6, I close with a summary, conclusions, and implications that define a new research agenda.

2 The linguistics of self-talk

In this section, I introduce some properties regarding the linguistics of self-talk that will serve as the backdrop relative to which questions of interest to theoretical linguists can be addressed. I start by discussing in more detail the phenomenon of self-talk and its significance (Section 2.1). Next, I discuss some of the methodological challenges that the exploration of self-talk brings with it (Section 2.2). Finally, I introduce several linguistic properties of self-talk that demonstrate that grammar is involved in regulating the language used in self-talk and, hence, that it constitutes a fruitful domain of study for theoretical linguists (Section 2.3).

2.1 The phenomenon and its significance

The significance of self-talk is perhaps best appreciated by considering the classic question as to whether human language is (primarily) a tool for thought or for

3 A similar distinction is made in Krifka (2023b), who refers to the two modes as *inner monologue* versus *inner dialogue* (see Section 5.1 for detailed discussion).

Table 1: Self-talk lies between thought and communication.

Thought	Self-talk	Communication
Private	Private	Social
Covert	Overt	Overt

communication (see Chomsky 2017 for the former view and Carruthers 2002; Jackendoff 2002; Levinson 2019 for the latter). Self-talk has the potential to be revealing in the context of this debate as – at least superficially – it appears to lie squarely between these two functions. On the one hand, self-talk has in common with thought that it is private (as opposed to the social dimension intrinsic in communication). On the other hand, self-talk has in common with communication that it is overt, or at least it can be (see below), whereas thought cannot (without turning into self-talk).⁴ This is summarized in Table 1.

Note for completeness that self-talk may remain covert. For example, Latinjak et al. (2023: 356) conceptualize self-talk as “verbalizations addressed to the self, overtly or covertly” (see also Hardy 2006; Theodorakis et al. 2000). The terminology surrounding the phenomenon of self-talk is messy. The term *self-talk* is sometimes reserved for overtly talking to oneself (alongside the terms *egocentric speech* and *private speech*) whereas other terms are used for covert self-talk, such as *inner voice*, *inner speech*, *verbal thoughts*, *covert speech*, *silent speech*, *verbal thinking*, *verbal mediation*, *inner monologue*, *inner dialogue*, *inner voice*, *articulatory imagery*, *voice imagery*, *speech imagery*, and *auditory verbal imagery* (see Alderson-Day and Fernyhough 2015; Nedergaard and Lupyan 2024 for an overview). When self-talk remains covert, it is impossible to distinguish from thought by the criteria listed in Table 1. According to Alderson-Day and Fernyhough (2015: 931), inner speech (a covert form of self-talk) “can be defined as the subjective experience of language in the absence of overt and audible articulation.”⁵

The terminological conventions, as well as the definitions of self-talk just reviewed, already point to different ways of viewing the nature of self-talk. Strikingly, this debate is not unlike the one regarding the function of language itself. Specifically, much of the discussion evolves around the question of whether self-talk should be viewed as a vehicle of thought or whether it is better analyzed as a way of communicating with oneself.

⁴ I use the terms overt and covert here, rather than, for example, silent versus audible, to reflect the fact that the medium for language is not necessarily auditory but can also be signed. Crucially, native speakers of sign language have been reported to engage in (signed) self-talk (Zimmermann and Brugger 2013).

⁵ In this paper I focus on overt self-talk, mainly for methodological reasons (see Section 2.2).

The first explicit mention of self-talk in Plato’s *Theaetetus* unites these views, suggesting that “the soul when thinking appears to me to be just talking – asking questions of herself and answering them, affirming and denying” (Plato 1970 (1892): 252). Accordingly, thinking is equated with self-talk, which, in turn, is a form of communication within oneself.

In contrast, the two pioneers of contemporary explorations of self-talk, Jean Piaget and Lev Vygotsky, held two opposing views regarding the relation between self-talk, thought, and communication. What they have in common is that both view self-talk as an important milestone in children’s cognitive development and, thus, as a window into the nature of cognition more generally.

Jean Piaget is generally considered the first psychologist who took an interest in self-talk (which he termed *egocentric speech*) and its role in cognitive development. He hypothesized that egocentric speech is a developmental step towards social speech (used in interaction with others). Accordingly, children start with their own thoughts, externalized in egocentric speech (labelled self-talk in Figure 1) and are only able to communicate them to others when they can take the perspective of others (Piaget 1923/1962).

A decade later, another psychologist, Lev Vygotsky – the first to conduct systematic experiments on children’s self-talk – develops a view diametrically opposed to that of Piaget. Specifically, according to Vygotsky (1934/1986), self-talk is rooted in communication with others (i.e., the social world of the child), which often serves the purpose of regulating the child’s behaviour. Over time, this is internalized, and the child uses (overt) self-talk for self-regulation. Eventually, this leads to complete internalization, i.e., covert self-talk and thought. Thus, for Vygotsky, overt self-talk bridges social speech and covert self-talk, which he viewed as cognitively more sophisticated than overt self-talk or communication with others. This is illustrated in Figure 2.

While in both views, self-talk constitutes an important developmental milestone, they differ as to whether it is viewed as an aid to externalizing thoughts to be communicated (Piaget) or as an aid to internalizing communicated content to arrive at the capacity of thought (Vygotsky). Even though both accounts focus on the developmental role of self-talk in children, these opposing views are also found in explorations of self-talk (overt and covert) in adults, who are known to use overt



Figure 1: The role of self-talk in cognitive development according to Piaget.



Figure 2: The role of self-talk in cognitive development according to Vygotsky.

self-talk while engaging in problem-solving and other activities (e.g., Duncan and Cheyne 2001; Duncan and Tarulli 2009; Winsler 2009). Some view self-talk (covert and overt) as a way to support the thinking process (Sokolov 1975), especially in demanding situations (Hardy 2006: 93), or even as a form of thinking itself. For example, according to Bunker et al. (1993: 226), “Anytime you think about something, you are in a sense talking to yourself.” This emphasizes the role of self-talk as a thinking tool (Geurts 2018; Lupyán and Swingley 2012). In contrast, others view self-talk as an intrinsically dialogical phenomenon through which “the individual interprets feelings and perceptions, regulates and changes evaluations and convictions, and gives him/herself instructions and reinforcement” (Diaz 1992; Hackfort and Schwenkmezger 1993: 355). This emphasizes the communicative character of self-talk (see Deamer 2021).

No matter whether self-talk is primarily viewed as a way of thinking or as a way of communicating with oneself, the fact remains that it seems to serve several functions in adults, including self-control, self-attention, self-regulation, self-motivation, self-critique (even self-denigration), developing metacognition and self-awareness, processing social situations, problem solving, among others (Alderson-Day and Fernyhough 2015; Brinthaup et al. 2015; Chella and Pipitone 2020; Kompa 2024). In the context of the present paper, the functions of self-talk are significant since there are studies which seek to correlate specific functions of self-talk with specific linguistic properties. For example, spontaneous and motivational self-talk have been argued to be characterized by the use of 1st person (e.g. *I’m actually good at this*, *I got this*.) whereas goal-directed and instructional self-talk by the use of 2nd person (e.g. *You are actually good at this*, *You need to mow the lawn this weekend!*) (Latinjak et al. 2023). Moreover, it has been shown that goal-directed self-talk is more efficient when formulated as a question (e.g., *Will I make it?*) or in the 1st person plural (e.g. *We will make it.*) (Senay et al. 2010; Son et al. 2011; Van Raalte et al. 2018).

There are three main messages to take away from this brief review:

- (8) Main lessons from the literature on self-talk
 - (i) Even though self-talk is a language-based phenomenon, the existing literature is almost exclusively restricted to the psychological and philosophical scholarship, with a linguistic perspective conspicuously absent.
 - (ii) Self-talk is not a unified phenomenon as it differs across various dimensions, including form (overt or covert), function (e.g., motivational, instructional, ...), and linguistic expression (e.g., 1st or 2nd person pronouns).

- (iii) Self-talk is situated in between thought and communication. As a limiting case, it thus has the potential to shed light on the nature of both.

The contribution of this paper is two-fold. First, by adding a linguistic perspective, we can shed light on some of the questions regarding the nature and function of self-talk. Second, by adding self-talk as a linguistic object of study, we can shed light on questions regarding the nature and function of language itself.

2.2 Methodologies

Having introduced the phenomenon of self-talk (and its kin), a few words are in order about methodological issues that arise when exploring (the linguistics of) self-talk.

As mentioned above, the first studies of self-talk involved children, who frequently talk to themselves and hence can be observed doing so. This is what both Piaget and Vygotsky did. However, when it comes to self-talk in adults, observation of self-talk in natural settings (and hence corpus data) is unavailable. This is because talking to oneself is associated with a social stigma – so much so that according to Goffman (1981: 81), “There are no circumstances in which we can say, ‘I’m sorry, I can’t come right now, I’m busy talking to myself.’” The reason for this stigma, according to Goffman (1981: 78), is that it violates the social agreement about the communicative function of speech. Accordingly, the apparent decrease of self-talk with age may be related to this social norm (Duncan and Tarulli 2009: 177). For all these reasons, the exploration of self-talk comes with methodological problems (Ariel 2022).

To overcome these difficulties, several methodologies have been employed. In psychological studies, self-talk is typically explored through self-reports. For example, Brinthaup et al. (2009) develop the Self-Talk Scale, which measures the frequency of self-talk and its correlation with particular functions of self-talk. However, this does not allow us to test for specific linguistic properties of the kind I wish to explore here, as it does not target specific linguistic forms.

One source for self-talk data comes from fictional contexts, literary and film alike (see Banfield 1982 for the significance of fiction for linguistic analysis). That is, in the context of movies or plays, characters are sometimes talking to themselves. In fact, there is a whole scholarship surrounding soliloquy in Shakespeare, for example (Murphy 2015). Some movies revolve around the premise that one of the characters can “hear” the thoughts (thus inner speech) of others, such as Mel Gibson’s character in *What Women Want* and Sookie Stackhouse in *True Blood*.

In the context of novels, too, self-talk in the form of an inner monologue is a common literary strategy through which a story unfolds. An early example of this is found in Arthur Schnitzler's *Leutnant Gustl*, whose inner speech has recently been analyzed in Krifka (2023b). While neither film, theatre, nor novels provide us with naturally occurring data, they at least give us some insights into the intuitions of their authors as to what constitutes well-formed self-talk.

This brings us to the main methodology used to collect data for the present paper: the use of native-speaker intuitions. Specifically, while data from fictional work gives us insight into what speakers might say in self-talk, they do not provide us with negative data or minimal pairs, as is the case for corpus data in general. Hence, linguists have long relied on native speaker judgements of constructed examples in the form of elicitation tasks, which have stood the test of time as a valid methodology (Arppe and Järviö 2007; Langsford et al. 2018; Schütze 2016, Sprouse and Almeida 2017). While classically, such elicitation tasks have been used to target language outside of interaction and, hence, without the specification of who is talking to who, it is safe to say that these tasks are typically not assumed to target language in the context of self-talk.

Using native speaker judgments to elicit the language used in self-talk thus relies on the premise that native speakers do indeed have judgements about what constitutes well-formed utterances in the context of self-talk. This assumption receives indirect support from the fact that speakers do have clear judgements about language in interaction, including differences pertaining to the identity of the interlocutor (Wiltschko 2021). For example, speakers have judgements about the use of formal and informal pronouns even in the absence of a relevant interlocutor. What makes the elicitation of language in interaction somewhat more involved than eliciting sentences in isolation is the fact that contexts must be explained, and consultants must pair a particular context with a particular utterance. For this reason, Wiltschko (2021) develops the conversation board methodology, which is based on Burton and Matthewson's (2015) storyboard elicitation task. This consists of presenting the consultant with cartoon-like pictures that minimally include a panel depicting the context of the utterance and a speech bubble for the interlocutor whose utterance is under investigation. The utterance may be presented within this speech bubble for a well-formedness judgement, or else the consultant may be asked to provide an appropriate utterance given the relevant context. In Goddard et al. (2022), this methodology was used for eliciting self-talk data where the context of self-talk was varied between thinking to oneself (depicted via a thought bubble), talking aloud to oneself (indicated by a speech bubble), talking to oneself in the mirror, and typical conversations with another interlocutor.

If not otherwise indicated, the data reported in this paper comes from simple, well-formedness judgements with native speakers. Judgements have been very clear and consistent across consultants. I also report on some of the data collected by Goddard et al. (2022), which were consistent with the findings obtained from simple, well-formedness judgements.

For completeness, note that by necessity, linguistic judgements can only be obtained for overt self-talk. This is because to judge the language of self-talk, it must be externalized (i.e., overt). These judgements might extend to silent self-talk, which has been reported to sometimes mirror overt self-talk and which has been shown to even activate brain areas that involve motor planning (e.g., Barch et al. 1999; Pratts et al. 2023; Yetkin et al. 1995). However, there are some versions of silent self-talk (i.e., inner speech) which cannot be observable, and which are known to differ linguistically in that they appear to be less articulated, condensed, and abstract (Alderson-Day and Fernyhough 2015: 942–943). Though this form of inner speech might have the potential to shed light on the nature of language and its relation to thought, I have nothing to say about this phenomenon. What follows is a discussion of overt self-talk only.

2.3 The linguistics of two modes of self-talk: an analytical challenge

As we have seen, the psychological literature recognizes differences in form, function, and linguistic expression, i.e., whether the person engaged in self-talk is using a 1st or 2nd person pronoun to refer to themselves. This linguistic difference will be at the core of the current exploration of self-talk. As a terminological convention, I use the terms *you-centered self-talk* and *I-centered self-talk* to distinguish between these two modes.⁶

A first question we may ask is if the difference between *I-centered* and *you-centered* self-talk correlates with any of the other dimensions of variation (i.e., form or function). The answer seems to be negative, as I now show. In terms of its form, both *I-centered* and *you-centered* self-talk can be used both overtly and covertly (Holmberg 2010). Next, consider the function of self-talk. Recall that it has been claimed that spontaneous and motivational self-talk are characterized by *I-centered* self-talk, whereas goal-directed and instructional self-talk are by

⁶ As Holmberg (2010) observes, 1st person plural pronouns can also be used in self-talk, as in *We can do this*. I exclude these cases from the current study. I hypothesize that they will pattern with *you-centered* self-talk in that they appear to be used when having a conversation with oneself. Though this hypothesis will have to be tested.

you-centered self-talk (Latinjak et al. 2023). If this is indeed the case, this suggests a correlation between the function of self-talk and its linguistic expression. However, this correlation does not seem to be categorical. While *I*-centered self-talk is certainly possible in spontaneous and motivational self-talk, it is not restricted to these functions. For example, the statement in (9), based on Holmberg (2010: 57) can hardly be classified as motivational, yet *I*-centered self-talk is possible, as in (9a). Moreover, while it might be classified as spontaneous (though it is unclear exactly what criteria one might apply), it still allows for *you*-centered self-talk, as in (9b).

- (9) a. I am an idiot.
b. You are an idiot.

Similarly, goal-oriented self-talk, too, allows for both *I*-centered and *you*-centered self-talk, as shown in (10).

- (10) a. I will go to the gym today, no matter what.
b. You will go to the gym today, no matter what.

Thus, the use of *I*-centered and *you*-centered self-talk does not categorically correlate with either its form or its functions.⁷ Nevertheless, there are crucial differences between *I*-centered and *you*-centered self-talk, as first observed in Holmberg (2010). These differences manifest themselves in different linguistic profiles, which in turn suggest that we are dealing with two different modes of self-talk (Ritter and Wiltschko 2021). In the remainder of this section, I review some of these distinctions. For each of the linguistic properties, I start by illustrating it based on a typical conversation, and then I show that it distinguishes between *you*-centered and *I*-centered self-talk.

The first phenomenon I consider is the use of a vocative nominal. Consider first the examples in (11) where Alaka is talking to Thea. Here, the use of the vocative is well-formed, regardless of whether Alaka is talking about Thea (using the 2nd person pronoun) or about himself (using the 1st person pronoun).

- (11) Alaka to Thea:
a. **Thea**, you're an idiot.
b. **Thea**, I'm an idiot.

7 The observations that led to the reported generalizations are likely frequency effects.

Now, consider what happens when Alaka is engaged in self-talk. In this case, referring to himself with a 2nd person pronoun is possible in the presence of a vocative, as in (12a). In contrast, referring to himself with a 1st person pronoun is not possible, as in (12b).

- (12) Alaka to himself:
- a. **Alaka, you're** an idiot.
 - b. ***Alaka, I'm** an idiot.

Krifka (2023b) finds support for this restriction on vocatives in the literary work he explores (Schnitzler's *Leutnant Gustl*). Specifically, he observes that passages that are characterized by *I*-centered self-talk switch to *you*-centered self-talk when a vocative is used. Significantly, there appears to be no instance of a vocative in *I*-centered self-talk.

The second phenomenon concerns imperatives. In a typical conversation, imperatives are characterized by a requirement for the subject to be equated with the addressee.⁸ This is illustrated in (13). While the subject of an imperative is often silent (indicated as *pro_{Adr}* in (13a)), it may remain overt. In this case, it may be realized as a 2nd person pronoun, as in (13b), or as a form of address (i.e., the addressee's name, as in (13c)). Crucially, 1st person pronouns are ungrammatical in this context, as shown in (13d).⁹

- (13) Alaka to Thea:
- a. *pro_{Adr}* stay positive!
 - b. You stay positive!
 - c. Thea stay positive!
 - d. *I stay positive!

Note further that even when the subject of an imperative remains silent, there is still evidence that it must be syntactically represented in the form of a silent pronoun, which refers to the addressee (*pro_{Adr}*). Specifically, when the direct object refers to the addressee, it has to be realized as a reflexive pronoun, as in (14a), whereas the use of a personal pronoun is ill-formed, as in (14b). This is because the silent *pro_{Adr}* binds the direct object, and hence, a reflexive pronoun is required as per binding theory. In contrast, when the direct object refers to the speaker, it is not bound by the silent

⁸ This definition excludes hortatives which involve speaker-oriented forms (either singular or plural). See van der Auwera et al. (2003) for a cross-linguistic perspective. It would be interesting to explore how hortatives behave in self-talk and if there are cross-linguistic differences depending on the exact properties of the imperative/hortative distinction.

⁹ For the purpose of this discussion, we abstract away from exhortatives, i.e., imperatives with 1st person plural subjects such as *Let's dance*.

*pro*_{Adr} subject. Hence, the reflexive 1st person pronoun is ungrammatical, as in (14c), whereas the personal pronoun is well-formed, as in (14d).¹⁰

- (14) Alaka to Thea:
- a. *pro*_{Adr} stop putting **yourself** down!
 - b. **pro*_{Adr} stop putting **you** down!
 - c. **pro*_{Adr} stop putting **myself** down!
 - d. *pro*_{Adr} stop putting **me** down!

In a typical conversation, the prohibition of the use of 1st person pronouns follows, of course, from the requirement that an imperative is necessarily addressed to someone else. But what happens in self-talk where the speaker is addressing themselves? Crucially, imperatives must still be realized with addressee denoting (2nd person) subjects. The data in (15) illustrate that in the context of self-talk, the same pattern holds as in a typical conversation: the subject may remain silent, as in (15a), and when it is overt, it must be realized as a 2nd person pronoun, as in (15b), or as a form of address (in this case the speaker's name), as in (15c). Crucially, the 1st person pronoun is ungrammatical, as shown in (15d).

- (15) Alaka to himself:
- a. *pro*_{Adr} stay positive!
 - b. You stay positive!
 - c. Alaka stay positive!
 - d. *I stay positive!

Thus, even though in self-talk the speaker and the addressee are the same person, and hence the speaker appears to address themselves, the 1st person pronoun is ruled out. In other words, imperatives are restricted to *you*-centered self-talk (Ritter and Wiltschko 2021).

This is further confirmed by imperatives with direct objects, where we witness the same generalization as with imperatives in typical conversations. Only the 2nd person reflexive is licit, as in (16a). As in typical conversations, the non-reflexive 2nd person pronoun is ungrammatical, as in (16b), suggesting that there is an

10 Vocatives present an interesting puzzle for the notion of binding. Specifically, the vocative in (i) is co-referent with the subject of the imperative, which in turn binds the reflexive pronoun.

(i) Thea_i, *pro*_{Adr-i} stop putting yourself_i / *herself_i down.

What is puzzling is that usually binding requires agreement in phi-features but the name used as a vocative is formally a third person nominal even though it refers to the addressee. And the fact that it refers to the addressee is consistent with the fact that the reflexive pronoun is obligatorily 2nd person. While a formalization of this property has to await another occasion, I submit that it may be accounted for on the assumption that phi-features are restricted to propositional structure and play no role in the interactional structure (see Ritter & Wiltschko 2024 for evidence to this effect).

addressee denoting *pro_{Adr}* in subject position. In contrast to imperatives in typical conversations, however, in self-talk, a 1st person pronoun is ungrammatical, no matter whether it is realized as a reflexive pronoun, as in (16c), or as a personal pronoun, as in (16d).

- (16) Alaka to himself:
- a. *pro_i* stop putting **yourself_i** down!
 - b. **pro_i* stop putting **you_i** down!
 - c. **pro_i* stop putting **myself_i** down!
 - d. **pro_i* stop putting **me_i** down!

Taken together, these facts demonstrate that imperatives are impossible in *I*-centered self-talk: the recipient of a command cannot be realized as a 1st person reflexive or personal pronoun.

The third property of self-talk I here discuss concerns verbs of cognition. Consider first the examples in (17), which involve a typical conversation. We can tell others something about our own mental state, as Alaka does in (17a). In contrast, we do not typically tell others about their current mental states. While there might be contexts in which this is possible, under normal circumstances, such utterances are deviant (hence, the example in (17b) is marked by a hashmark #).

- (17) Alaka to Thea
- a. I can't believe my luck.
 - b. # You can't believe your luck.

The deviance of (17b) follows from the conditions of use for assertions. Asserting a proposition *p* is felicitous when two conditions hold: (i) the speaker is certain about the truth of *p* and (ii) the speaker assumes that the addressee does not know *p*. Since someone's mental state is accessible to them and only to them, one cannot felicitously tell someone else about their mental state. Thus, even if Alaka is certain about Thea's mental state (because Thea has a certain expression on her face), he could still not tell her about it because Thea would already know. Hence, (17b) is ill-formed.

Instead, what Alaka might felicitously say in contexts where he observes an expression of disbelief on Thea's face is an assertion of having evidence for her mental state, as in (18a). This is well-formed because Alaka has direct access to how Thea looks, and Thea may not know that she does look like she can't believe her luck. Thus, in this context, the utterance in (18a) follows the felicity conditions of assertion. And now the converse holds for (18b). Under normal circumstances, a speaker has access to their own mental state, and hence, they would not resort to talking about overt evidence for their mental state. Additionally, such evidence would be accessible to an addressee, and hence, telling them what the evidence suggests again violates the conditions for assertion. Hence, (18b) is ill-formed.

- (18) Alaka to Thea.
- a. It looks like you can't believe your luck.
 - b. # It looks like I can't believe my luck.

The same holds for other ways of stating one's evidence for another person's mental state, as shown in (19).

- (19) Alaka to Thea
- a. Apparently, ...
 - b. You are acting like...
 - c. It seems to me like...
...you can't believe your luck.
... I can't believe my luck.

Now, consider what happens in self-talk. As Holmberg (2010) observes, verbs of cognition and certain experienter verbs are not allowed in *you*-centered self-talk, whereas they are in *I*-centered self-talk (adapted from Holmberg 2010: 59f.).¹¹

- (20) Alaka to himself
- a. # You can't believe your luck.
 - b. I can't believe my luck.
- (21) Alaka to himself
- a. # You can't take this anymore.¹²
 - b. I can't take this anymore.

About this contrast, Holmberg (2010: 60) states that “*you* can't refer to the self as an experienter of feelings or holder of intentions or plans [...]”. He proposes that the referent of *you* in self-talk is a “mindless self” and hence cannot serve as the subject of a verb of cognition. While Holmberg recognizes that the same pattern holds in typical conversations, he suggests that the restrictions, though seemingly identical, have a different explanation. In a typical conversation, *you* refers to the addressee whose mind is inaccessible to the speaker. Clearly, in self-talk, the addressee's mind is accessible, and hence Holmberg (2010: 60f.) suggests that *you* in self-talk refers to a mindless self.

¹¹ I here follow Holmberg's description of these examples as involving verbs of cognition. A systematic exploration of what types of predicates fall under this umbrella is required but goes beyond the scope of the present paper.

¹² The ill-formedness of this example is restricted to the epistemic reading of *can* (akin to *you are not able to*). The sentence is well-formed under a deontic reading (akin to *you should not*).

This explanation runs into problems, however, when we consider the examples in (22). As in a typical conversation, the restriction against *you* being used as the subject of *believe* disappears when embedded under “*It looks like...*”, as shown in (22a). In contrast, in this context, *I* cannot be used as the subject of *believe*, as shown in (22b).

- (22) Alaka to himself
- a. It looks like you can’t believe your luck.

b. # It looks like I can’t believe my luck.
- (23) Alaka to himself
- a. Apparently, ...

b. You are acting like...

c. It seems to me like...

...you can’t believe your luck.

... I can’t believe my luck.

The conclusion we can draw from these data is that *you* in self-talk cannot be analyzed as referring to a mindless self, contrary to Holmberg (2010). From a theoretical perspective, this does not seem to be surprising. It is not clear why the addressee-denoting second person pronoun should have different referential properties in self-talk. Rather, *you* in self-talk behaves just like *you* in a typical conversation: it appears to refer to an inaccessible mind. But this is, of course, counterintuitive. A person engaged in self-talk has access to their own mind regardless of whether they use *you*-centered or *I*-centered self-talk.

Thus, the properties of utterances involving verbs of cognition highlight the analytical challenge the two modes of self-talk present us with. That is, we have now seen three linguistic properties that distinguish *I*-centered from *you*-centered self-talk. While *I*-centered self-talk is incompatible with vocatives and imperatives, *you*-centered self-talk does not support talking about one’s mental state. Thus, across all three properties, *you*-centered self-talk behaves no differently from typical conversations. This is summarized in Table 2.

Table 2: Empirical differences among modes of talking.

	<i>I</i> -centered self-talk	<i>You</i> -centered self-talk	Typical conversation
Vocatives	x	✓	✓
Imperatives	x	✓	✓
Verbs of cognition	✓	x	x

Now, consider why the properties of self-talk constitute a conundrum that needs to be addressed. It may be tempting to assume that the difference between self-talk and typical conversations is pragmatic in nature. After all, these two modes of talking are defined by contextual variables, namely, who is talking to whom. At first glance, this invites a pragmatic solution. However, such a solution runs into non-trivial problems in the face of the two modes of self-talk.

Consider first what a pragmatic account for the ban on imperatives in *I*-centered self-talk might look like. I assume that speech acts realized as imperatives are associated with two conditions of use when used canonically: (i) the speaker wants the addressee to act in a certain way and (ii) the addressee would not act in this way if the speaker were not to utter the imperative. Arguably, imperatives are ill-formed in *I*-centered self-talk for the same reason as imperatives cannot have a 1st person singular subject. A speaker cannot command themselves to initiate an action, and this is true no matter if the speaker talks to themselves or to someone else (in which case they would indicate to their addressee that they are commanding themselves). While this might be a plausible account for the ill-formedness of imperatives in *I*-centered self-talk, it cannot account for their well-formedness in *you*-centered self-talk. This is because the context that characterizes self-talk is identical, no matter whether it is *I*-centered or *you*-centered. In both types of self-talk, one is giving a command to oneself, and hence, they should equally be ruled out under the pragmatic account sketched above.

Similar considerations hold for the use of vocatives in self-talk. That is, the function of a vocative is to either get the addressee's attention or to maintain it (Zwicky 1974). On a pragmatic account, one might reasonably propose that one need not and, therefore, cannot get one's own attention. This would explain the ill-formedness of vocatives in *I*-centered self-talk. However, this explanation cannot straightforwardly account for the well-formedness of vocatives in *you*-centered self-talk. Again, from a pragmatic point of view, both types of self-talk are characterized by the same context (a person talking to themselves) no matter which pronoun they use to refer to themselves. Thus, a pragmatic account might succeed in accounting for the ill-formedness of imperatives and vocatives in *I*-centered self-talk, but it cannot account for their well-formedness in *you*-centered self-talk.

As for the properties of cognitive predicates in the context of self-talk, here, a pragmatic account cannot straightforwardly explain the ill-formedness of *you*-centered self-talk. That is, if the reason one cannot use *you* as a subject of a cognitive predicate is that one does not have access to somebody else's mind, then why does the same hold in self-talk? So why is *you*-centered self-talk ruled out when talking about one's mental state? And why does it matter whether the person engaging in self-talk refers to themselves with a 1st or 2nd person pronoun?

In a nutshell, the analytical challenge surrounding self-talk presents itself differently for *I*-centered and *you*-centered self-talk. In both cases, the context is such that a speaker is talking to themselves, and thus, the speaker simultaneously serves as their addressee. Yet, the linguistic properties of *I*-centered self-talk suggest that there is no addressee, ignoring the real-world fact that there is, even if it is identical to the speaker. In contrast, the linguistic properties of *you*-centered self-talk suggest that there is an addressee, but what is ignored in this case is the real-world fact that the addressee is identical to the speaker; it has the same properties as in a typical conversation in which the addressee is different from the speaker. This is something a pragmatic account cannot explain because we would not expect that real-world knowledge be ignored. Rather, what the linguistic profile of self-talk suggests is that the properties are grammatical in nature. That is, it is the hallmark of grammar that it may remain insensitive to real-world knowledge. For example, while grammatical gender may be rooted in real-world properties, categorizing a noun as masculine or feminine may remain insensitive to real-world knowledge. What the linguistic properties of self-talk thus reveal is that the addressee role must be part of the grammatical representation of an utterance. As mentioned above, there is growing consensus that this is indeed the case, based on several linguistic phenomena that demonstrate a sensitivity to the presence of an addressee. In the next section, I shall briefly review some of these proposals, which will then allow us to evaluate them in a novel way based on how they fare in deriving the linguistic properties of self-talk.

3 Modelling the grammar of language in interaction

The goal of this section is to introduce current grammatical models of language in interaction, which often include a representation of the speaker and the addressee, albeit in different ways. On the one hand, this will set the stage for exploring the properties of self-talk in a theoretically informed way and on the other hand, we shall see that self-talk may be used as a window into the proper modelling of language in interaction.

Modern formal linguistics has long concerned itself with sentences in isolation, often considered to be the pure expression of thought. In this way, modern linguistics followed in the footsteps of the ancient grammarians. However, Austin's and Searle's speech act theory has changed this to some degree. Specifically, according to Austin (1962), we do things when we say things, especially in performative acts, which have the power to change the world. Austin suggests that this performative dimension is not restricted to explicitly performative utterances (*I hereby order you to leave*). The

performative aspect may remain implicit (as in *Leave!*). This insight led Ross (1970) to argue two things (see also Sadock 1969a, b). First, Ross argues that basic declarative sentences are also implicitly performative. Namely, they are communicating the content of their utterance (which may correspond to their thought) to an addressee. Second, Ross argues that this communicative act itself has a grammatical representation. Specifically, Ross suggests that any declarative clause is embedded in a *speech act structure* encoding the speaker (in the form of a 1st person pronoun), the addressee (in the form of a 2nd person pronoun), and a verb of communication, thus forming the typical clause-structure of the time, as illustrated in (24). Notably, the speech act structure is assumed to undergo deletion and hence is silent, as indicated by the strikethrough in (24).

(24) $\text{st}_{\text{NP}} \text{[I]} \text{VP} \text{[tell]} \text{NP} \text{[you]} \text{S} [\text{p-structure}]$

The presence of the speech act structure has the effect of encoding the illocutionary force syntactically, such that a declarative clause is interpreted as the speaker telling the addressee the propositional content. On this view, even a declarative clause is performative in the sense that the speaker performs the act of telling their addressee the propositional content – hence, this analysis is known as the *performative hypothesis*. While it was quickly dismissed on theoretical and empirical grounds (Anderson 1971; Fraser 1974; Leech 1976; Mittwoch 1976, 1977), the idea of a speech act structure was revived in several different ways, which Wiltschko and Heim (2016) refer to as *neo-performative hypotheses*. While these analyses differ both in terms of analytical details and empirical coverage, they have at least two aspects in common, which differentiate them from the Ross/Sadock type performative hypothesis. First, speech act structure is viewed as being comprised of projections of functional categories (rather than projections of lexical categories, i.e., NP and VP). Second, this speech act structure is not assumed to undergo deletion but rather consists of abstract functional heads (which may remain silent), but which may be spelled out by various units of language used in typical conversations (such as vocatives and certain sentence-final particles, for example).

In their seminal paper, Speas and Tenny (2003) explicitly resurrect the Ross/Sadock idea of a dedicated speech act phrase (saP), which dominates the propositional structure. Other scholars had already proposed ideas along those lines (Ambar 1999; Cinque 1999; Etxepare 1997; Rizzi 1997), though Speas and Tenny (2003) are the first to conceptualize this structure as a speech act structure rather than as being part of an articulated CP. Specifically, Speas and Tenny propose that saP introduces the speaker and the addressee in the form of speech act *roles* rather than 1st and 2nd person pronouns, as in Ross (1970). The essence of their proposal is schematized in (25) (adapted from Speas and Tenny 2003: 320). (Note that, for ease of exposition and comparison, I have amended their original structure by replacing their *Utterance*

content with *p* (for propositional content) and their *Hearer* with *Adr*. In the remainder of this paper, I shall follow this adapted notation).

(25) $[_{sap} \text{ Spkr } [sa] [_{sap} p [_{sa} [sa] \text{ Adr}]]]$

Significantly, the Ross/Sadock insight is preserved in this proposal insofar as the speaker and addressee roles are analyzed akin to thematic roles in the event structure encoded by an articulated vP. Thus, according to (25), the speech act structure parallels a double object construction where the speaker (i.e., the agent) “gives” the utterance content (i.e., the theme) to the hearer (i.e., the receiver). Among their arguments for the postulation of this speech act structure (and its parallelism to the vP-level event structure) is the cross-linguistic generalization that there is a limited set of clause types (Sadock and Zwicky 1985) and speech act roles just as there are a limited number of (grammaticalized) event types and event roles (Hale and Keyser 2002; though see Gärtner and Steinbach 2006 for some critical remarks). Moreover, their analysis allows them to address the longstanding problem of the mapping between clause type and speech act type. Specifically, they argue that the structure in (25), repeated below as (26a), derives the fact that declaratives are interpreted as assertions. As for interrogatives being interpreted as questions, Speas and Tenny propose that this derives from the addressee role being moved to a position higher than the utterance content, as in (26b) (adapted from Speas and Tenny 2003: 320).

- (26) a. Declarative: $[_{sap} \text{ Spkr } [sa] [_{sap} p [_{sa} [sa] \text{ Adr}]]]$
 b. Interrogative: $[_{sap} \text{ Spkr } [sa] [_{sap} \text{ Adr } [p [_{sa} [sa] \text{ Adr}]]]]]$

The same movement of the addressee role is postulated for imperatives, which are interpreted as directives. This flip results in an interpretation whereby the addressee is given epistemic priority over the (interrogated or directed) utterance content. The difference between interrogatives and imperatives, according to Speas and Tenny, reduces to the finiteness of the utterance content ([+finite] for interrogatives and [-finite] for imperatives).

Since Speas and Tenny (2003), the literature on the syntacticization of speech act structure has grown significantly, and the empirical domains it serves to cover have expanded (see Wiltschko 2021 for a detailed overview). One of the striking pieces of evidence, already discussed in Ross (1970) – is the fact that the speech act roles may trigger agreement, which has been shown to be syntactic in nature (Haddican 2015, 2018; Oyharçabal 1993; Miyagawa 2022; Zu 2015, 2017). While details and empirical coverage differ across the various analyses, I here classify all approaches according to which the speech act structure contains a speaker role and an addressee role, with the former realized higher than the latter, as *neo-performative* analyses. This is because most of these accounts can be considered an updated version of the

Ross/Sadock analysis and are thus explicitly or implicitly inspired by the speech act theory of Austin and Searle.

There are, however, also analyses of the structure at the very top of the clause which take inspiration from some of the descendants of speech act theory. One is that of Krifka (2013, 2015), which is based on a dynamic semantics of speech acts and common ground updates. It incorporates insights of Krifka's *commitment space semantics* into the syntax at the top. I thus refer to this approach as *commitment-based* analyses (see also Miyagawa and Hill 2023). In its most articulated version, Krifka (2023a) assumes that the propositional structure is embedded in an articulated structure consisting of a JudgeP, a CommitP, and an ActP, as in (27).

(27) [ActP Act [CommitP Commit [JudgeP Judge [TP p]]]]

Much of Krifka's motivation for this structure comes from the interpretation of different illocutionary forces, different types of questions, as well as the interpretation of adverbial modifiers. JudgeP is responsible for introducing a private perspective, CommitP is responsible for introducing a public commitment to the truth of the proposition, and ActP is responsible for the illocutionary force of the utterance (i.e., assertion, marked as • in (28) (Krifka 2023b: 350) versus question, marked as ? in (29) (Krifka 2023b: 351)). Krifka (2023a) further incorporates various contextual parameters, including s(peaker), a(ddressee), j(udge), and c(ommitter). In a typical assertion, the committer is also the judge (marked as j:=c in (28)), and the speaker is the committer (marked as c:=s in (28)).

(28) Assertion:

[ActP • c:=s[CommitP c commits to j:=c[JudgeP j judges [TP ...]^{j,s,a}]^{s,a}]^{s,a}]

In contrast, in a typical question, it is the addressee who serves as the committer (marked as c:=a in (29)).

(29) Question:

[ActP ? c:=a[CommitP c commits to j:=c[JudgeP j judges [TP ...]^{j,s,a}]^{s,a}]^{s,a}]

Thus, Krifka's commitment-based analysis differs from the neo-performative one in that the speaker and addressee roles are not introduced by dedicated functional projections but as parameters in the semantic representation. Instead, the functional projections introduce various roles that may be realized by either the speaker or the addressee. These roles are defined by the relation the interlocutors have to the propositional content.

Another approach to analyzing the structure at the very top takes inspiration from approaches to language that focus on *conversational interaction*. That is, classic speech act theory, and many of its descendants, are concerned with a proper understanding of meaning in language viewed as an *act* of doing something with

words. The role of the addressee in the construction of meaning is hardly addressed. While it is acknowledged that in addition to the *locution* (what is being said) and the *illocution* (what is intended by the speaker), there is also a *perlocution* (what is effected in the addressee), the latter is hardly ever explored (see Gaines 1979; Lee 1974; Marcu 2000; Weigand 2010 for some exceptions). This is in stark contrast with approaches that focus on the logic of conversational interaction, such as *conversation analysis* (Sacks et al. 1974) and *grounding theory* (Clark 1992, 1996), which highlight the importance of the addressee. Conversation analysis focuses on the regularities of turn-taking while grounding theory focuses on the regularities and interactional nature of common ground construction. It is these two functions of language that Wiltschko (2021) incorporates into the grammatical structure that defines an utterance. Hence, I refer to this type of analysis as *interaction-based*. Specifically, as shown in (30), the propositional structure is immediately dominated by $\text{Ground}_{\text{Spkr}}\text{P}$, which is dedicated to marking the status of the propositional content relative to the speaker's epistemic state (i.e., the speaker's ground). $\text{Ground}_{\text{Spkr}}\text{P}$ is further dominated by $\text{Ground}_{\text{Adr}}\text{P}$, which is dedicated to marking the status of the propositional content relative to (the speaker's assumptions regarding) the addressee's epistemic state (i.e., the addressee's ground). The topmost category ($\text{Resp}(\text{onse})$) is dedicated to marking the status of the utterance for the purpose of turn-taking. Specifically, RespP can be self-oriented, deriving a reaction turn via marking an utterance as belonging in one's own (i.e., the speaker's) response set, as in (30a), but RespP can also be other-oriented, deriving an initiating turn via marking an utterance as belonging in the addressee's response set, thus requesting a response, as in (30b).

- (30) a. Reaction move:
 $[\text{RespP } \text{Resp-set}_{\text{Self}} [\text{Ground}_{\text{Adr}}\text{P } \text{Ground-Adr } [\text{Ground}_{\text{Spkr}}\text{P } \text{Ground-Spkr } [p]]]]$
 b. Initiating move:
 $[\text{RespP } \text{Resp-set}_{\text{Other}} [\text{Ground}_{\text{Adr}}\text{P } \text{Ground-Adr } [\text{Ground}_{\text{Spkr}}\text{P } \text{Ground-Spkr } [p]]]]$

The empirical motivation for this interactional structure comes from the grammatical properties of confirmationals (sentence-final particles used to request confirmation), response markers, as well as some intonational contours.

Unlike the neo-performative analyses, but like Krifka's commitment-based analysis, Wiltschko's interaction-based analysis does not include the speaker and addressee role via dedicated functional projections. Rather, these roles are encoded indirectly via introducing them as *ground-holders* indexed to the speaker and the addressee, respectively. In addition, the response set (which can be viewed as the grammatical representation of the *Table* in the sense of Farkas and Bruce (2010)) is indexed either to the *SELF* (hence the speaker) or to an *OTHER* (hence the addressee). Thus, Wiltschko's interactional structure differs from the other two approaches in that it does not contain a dedicated (Speech) Act Phrase, the conceptual argument

Table 3: Three approaches toward the syntax at the top.

	Neo-performative	Commitment-based	Interaction-based
Hierarchy:	S > A	Act > Commit > Judge	Resp > Ground _{Adr} > Ground _{Spkr}
Regulates:	Speech acts	Speech acts (dynamic)	Interaction
Roles:	Speaker	Judge	Ground holders
	Addressee	Committer	Turn-holders

being that speech acts are constructed and hence cannot define grammatical categories (Heim 2019; Heim and Wiltschko 2020; Wiltschko 2021). It further differs from the neo-performative approach in that the addressee-oriented category Ground_{Adr} dominates the speaker-oriented one Ground_{Spkr}, while the neo-performative analysis adopts the Ross/Sadock analysis according to which the speaker is higher than the addressee.

In sum, we have now seen three different approaches towards the syntax at the top. What they all have in common is the assumption that the propositional structure is embedded in structure which encodes information that has long been considered to be pragmatic in nature. Though, as we have seen, there are significant differences across the three types of approaches. As summarized in Table 3, these differences concern the hierarchy of functional categories postulated, the question as to what this structure is meant to regulate, and the pragmatic roles it is assumed to introduce.

One of the goals of this paper is to introduce self-talk as a litmus test to evaluate the empirical adequacy of these approaches.

4 An interaction-based analysis of self-talk

In this section, I introduce the interaction-based analysis of self-talk, first proposed by Ritter and Wiltschko (2021). To the best of my knowledge, this is the first explicit grammatical analysis of self-talk. Specifically, Ritter and Wiltschko propose that the empirical differences between *I*-centered and *you*-centered self-talk, discussed in Section 2.3, are structurally conditioned. Accordingly, *I*-centered self-talk has a speaker-oriented grounding phrase but lacks the addressee-oriented one, as in (31a). In contrast, *you*-centered self-talk is characterized by having both positions available, as in (31b). In addition, Ritter and Wiltschko propose that both types of self-talk are characterized by the absence of the topmost interactional category (RespP), and hence, they differ from the interactional structure available in a typical conversation, as in (31c).

- (31) The structural deficiency of self-talk
- a. [Ground-Spkr [p]] *I*-centered self-talk
 - b. [Ground-Adr [Ground-Spkr [p]]] *you*-centered self-talk
 - c. [RespP [Ground-Adr [Ground-Spkr [p]]]] typical conversation

Thus, according to this analysis, self-talk is structurally deficient compared to typical conversations, and *I*-centered self-talk is structurally deficient compared to *you*-centered self-talk.

The goal of this section is to present their analysis in more detail and show how it derives the empirical differences among the three modes of talking. I also add additional empirical evidence based on discourse markers. Furthermore, I discuss the implications of the interaction-based analysis of self-talk for our understanding of the phenomenon more generally.

4.1 The absence of an addressee-oriented layer in *I*-centered self-talk

Recall from Section 2.3 that *I*-centered self-talk is characterized by two restrictions that differentiate it from *you*-centered self-talk: it does not license the use of vocative nominals, nor does it allow imperatives. The relevant examples are repeated below for convenience.

- (12) Alaka to himself:
- a. **Alaka**, you're an idiot.
 - b. ***Alaka**, I'm an idiot.
- (16) Alaka to himself:
- a. *pro*_i stop putting **yourself**_i down!
 - b. * *pro*_i stop putting **you**_i down!
 - c. * *pro*_i stop putting **myself**_i down!
 - d. * *pro*_i stop putting **me**_i down!

According to the proposal in (31), *I*-centered self-talk is structurally deficient as compared to *you*-centered self-talk in that it lacks the addressee-oriented grounding layer. This derives the ban on vocatives and imperatives in *I*-centered self-talk as follows.

On independent grounds, Ritter and Wiltschko (2020) have argued that vocatives occupy the addressee-oriented grounding layer (i.e., SpecGround_{Adr}P). This analysis captures the fact that vocative nominals name the addressee and serve various functions, such as getting the addressee's attention or alerting the addressee that the content of the utterance is particularly relevant to them (Zwicky 1974). It thus follows

that vocatives are not licensed in *I*-centered self-talk as it lacks the relevant interactional position. Conversely, in *you*-centered self-talk, $\text{Ground}_{\text{Adr}}$ is available, and hence vocatives are well-formed. This is schematized in (32), where the asterisk indicates that the vocative has no available position to be licensed.

- (32) a. ***Vocative** $[\text{Ground-Spkr} [\text{p}]]$ *I*-centered self-talk talk
 b. $[\text{Ground-Adr}$ **Vocative** $[\text{Ground-Spkr} [\text{p}]]]$ *you*-centered self-talk

As for the impossibility of imperatives in *I*-centered self-talk, again, it follows from the absence of $\text{Ground}_{\text{Adr}}\text{P}$. Suppose that the null subjects (pro_{Adr}) of imperatives must be coindexed with the addressee in $\text{SpecGround}_{\text{Adr}}\text{P}$ (see Ritter 2024 for a suggestion along these lines). In the absence of $\text{SpecGround}_{\text{Adr}}\text{P}$, there is no antecedent for the subject of the imperative. It thus follows that imperatives are not licensed in *I*-centered self-talk as it lacks the relevant interactional position that would provide the antecedent for *pro*. Conversely, in *you*-centered self-talk, $\text{Ground}_{\text{Adr}}$ is available, and hence imperatives are well-formed. This is schematized in (33), where the asterisk indicates that the addressee has no available position to be licensed.

- (33) a. ***Adr** $[\text{Ground-Spkr} [\text{CP pro} \dots]]$ *I*-centered self-talk
 b. $[\text{Ground-Adr}$ **Adr** $[\text{Ground-Spkr} [\text{CP pro} \dots]]]$ *you*-centered self-talk

Another difference that sets apart *I*-centered from *you*-centered self-talk concerns the use of discourse markers, i.e., units of language which serve to regulate the interaction and thus many of them are addressee-oriented. They are precisely the empirical domain that motivated the interactional spine hypothesis. That is, according to Wiltschko (2021), sentence-peripheral discourse markers are directly associated with the interactional spine. Hence, it is predicted that discourse markers which are licensed in $\text{Ground}_{\text{Adr}}$ will not be available in *I*-centered self-talk. This prediction is borne out.

First, consider *huh*, which can be used as a confirmational. As shown in (34), in a typical conversation, *huh* can be used by the speaker to request confirmation for their assumption that the addressee holds the belief expressed in the proposition. In (34a), the belief to be confirmed is that the addressee should read *Moby Dick*, and in (34b), the belief to be confirmed is that the speaker should read it. In both cases, the speaker wants confirmation that the addressee believes the proposition to be true.

- (34) Alaka to Thea:
 a. **You** should read *Moby Dick*, **huh**?
 b. **I** should read *Moby Dick*, **huh**?

That *huh* is indeed addressee-oriented is supported by the fact that it cannot be used to confirm a belief the speaker firmly holds. This is shown by the contrast in (35). One

does not have direct access to another person's mental state, and hence, a speaker can use *huh* only to confirm their assumptions about someone else's mental state, as in (35a). In contrast, if the propositional content is such that the speaker has direct epistemic access, as when it concerns their own mental state, then the use of *huh* is ruled out, as in (35b).

- (35) Alaka to Thea:
 a. You like *Moby Dick*, **huh**?
 b. *I like *Moby Dick*, **huh**?¹³

Contrasts of this kind led Wiltschko (2021) to assume that *huh* must be interpreted in Ground_{Adr}P – it is always about the addressee's epistemic state. This is schematized in (36).

- (36) [Ground-Adr **huh** [Ground-Spkr [p]]]

Now, consider what happens in self-talk. As shown in (37), *huh* is possible in *you*-centered self-talk but not in *I*-centered self-talk (cf. Ritter and Wiltschko 2021).

- (37) Alaka to himself:
 a. You should read *Moby Dick*, **huh**?
 b. *I should read *Moby Dick*, **huh**?

This follows from the analysis of self-talk. Since *I*-centered self-talk lacks Ground_{Adr}P, *huh* is not licensed (as indicated by the asterisk in (38a)). This contrasts with *you*-centered self-talk, which contains Ground_{Adr}P and hence, *huh* is well-formed, as in (38b).

- (38) a. ***huh** [Ground-Spkr [p]] *I*-centered self-talk
 b. [Ground-Adr **huh** [Ground-Spkr [p]]] *you*-centered self-talk

Note that a purely pragmatic analysis would not be able to account for the contrast in (37). That is, one might argue that (37b) is ruled out because one cannot request confirmation from oneself. However, the same logic would equally apply to *you*-centered self-talk, as here, too, one is requesting confirmation from oneself. In the analysis in (38), however, the contrast in (37) is grammatically conditioned. Hence, real-world knowledge cannot interfere with the grammatical restriction on *I*-centered self-talk. In the absence of Ground_{Adr}P, there is no grammatically licensed

¹³ There is a well-formed reading available for (35a), which is one where Alaka wants to confirm that Thea thinks that he likes *Moby Dick*. That is, Alaka might have heard that Thea told someone else about it. The well-formedness of (35b) in this context further supports Wiltschko's (2021) analysis of *huh* according to which it is always used to confirm the speaker's assumption about their addressee's belief.

position for an addressee, even if the speaker knows that they are addressing themselves.

Next, consider the use of the discourse particles *doch* and *nämlich*, which are sensitive to the addressee's epistemic state. As we have seen in Section 1, *doch* is used when the speaker assumes that the addressee knows (or should know) the propositional content of their utterance. The sensitivity of *doch* to the epistemic state of the addressee can be analyzed by assuming that it associates with $\text{Ground}_{\text{Adr}}$ (either via agree or covert movement; see Thoma 2016; Wiltschko 2024b). Hence, we expect that *doch* is sensitive to the mode of self-talk. It should be possible to use *doch* in *you*-centered self-talk but not in *I*-centered self-talk.¹⁴ This prediction is borne out, as shown in (39).

- (39) Alaka is trying on a pair of his jeans, which he likes but which have been too small for him for a while now. As he is struggling to put them on, he says to himself:

- a. Die pass-n da **doch** ned
 DET.PL suit-3PL 2SG.DAT DOCH NEG
 'These don't fit you.' [and you should know that]
- b. *Die pass-n ma **doch** ned
 DET.PL suit-3PL 1SG.DAT DOCH NEG
 'These don't fit me.' [and I should know that]

Now consider *nämlich*, which is used when the speaker assumes that the addressee does not (or cannot) know the propositional content. Thus, *nämlich* is ruled out in self-talk, no matter whether it is *you*-centered, as in (40a) or *I*-centered, as in (40b).

- (40) Alaka wants to buy a pair of jeans and is trying on a pair that he likes. Even though they are the size he normally wears, they turn out to be too small. As he is struggling to put them on, he says to himself:

- a. *Die pass-n da **nämlich** ned
 DET.PL suit-3PL 2SG.DAT NÄMLICH NEG
 'These don't fit you.' [and you do not know that]
- b. *Die pass-n ma **nämlich** ned
 DET.PL suit-3PL 1SG.DAT NÄMLICH NEG
 'These don't fit me.' [and I do not know that]

¹⁴ The discussion below is restricted to unstressed *doch*. As is well-known, stressed *DOCH* has different conditions of use (see Egg and Zimmermann 2011), which are compatible with *I*-centered self-talk.

The ungrammaticality of *nämlich* in self-talk is due to the simple fact that it cannot be the case that a person engaged in self-talk will assume that they do not (or cannot) know the propositional content that they are uttering. In this way, the relation of the person engaged in self-talk to themselves is akin to the relation between two close acquaintances (see example (5), Section 1).

4.2 The addressee in you-centered self-talk

In this section, we turn to the restriction on *you*-centered self-talk discussed in Section 2.3, namely the ban on verbs of cognition. I show how it can be analyzed and argue that it provides additional evidence for a grammatical approach (Section 4.2.1). Moreover, I discuss a prediction of the present analysis, namely that we expect two different types of *you*-centered self-talk. That is, given that *you*-centered self-talk is analyzed as a conversation one has with oneself, we expect that the actual self may hold either of the two available interactional roles. That is, a person engaged in self-talk may take on the role of the speaker addressing a mirror image of themselves (for example), or else, they may take on the role of the addressee being addressed by an imaginary speaker. I present evidence that this prediction is indeed borne out (4.2.2).

4.2.1 The addressee role is always grammatically construed as another mind

Recall that verbs of cognition are impossible in *you*-centered self-talk, just as they are in typical conversations. This was illustrated with the data in (20), repeated below for convenience.

- (20) Alaka to himself:
- a. # You can't believe your luck.
 - b. I can't believe my luck.

As discussed in Section 2.3, Holmberg (2010: 60) attributes the impossibility of the use of verbs of cognition in *you*-centered self-talk to the assumption that the addressee in self-talk is a “mindless self.” I agree that *you* cannot refer to the self as a holder of thoughts or beliefs, but not because it is a mindless self. Rather, the use of *you* signals the presence of an addressee whose mind is not accessible to the speaker. Thus, (20a) is ill-formed regardless of whether it is uttered in self-talk or in a typical conversation. In turn, this is consistent with the proposal according to which *you*-centered self-talk (unlike *I*-centered self-talk) is characterized by the presence of Ground_{Adr}. The grammatical representation of the holder of addressee-ground is, by hypothesis, construed as a mind inaccessible to the speaker. It is the mind of another who the speaker has no access to. In other words, Ground_{Adr} is not a direct representation of

the addressee's knowledge state (their ground) but is instead a representation of the speaker's assumptions about the addressee's knowledge state. This accounts for the fact that, unlike the constraints on *I*-centered self-talk, the constraint illustrated in (20) is not restricted to *you*-centered self-talk but is a general constraint on interactions between speaker and addressee. Thus, the use of *you* signals the presence of an inaccessible mind, irrespective of who that mind belongs to. Real-world knowledge (such as the fact that in self-talk, the speaker and the addressee are the same person) is inaccessible to grammar. In other words, in *you*-centered self-talk (just as in a typical conversation), the speaker treats the addressee as an inaccessible mind belonging to an *OTHER* and not as a mindless individual. This is illustrated in (41).

- (41) a. [Ground-Spkr SELF [p]] *I*-centered self-talk
 b. [Ground-Adr OTHER [Ground-Spkr SELF [p]]] *you*-centered self-talk

That real-world knowledge cannot override grammatically conditioned interpretation is well-established. In this way, the impossibility of verbs of cognition in *you*-centered self-talk provides support for the claim that the addressee role does have a grammatical representation.

4.2.2 Two modes of *you*-centered self-talk

We now turn to a prediction of the analysis in (31), according to which *you*-centered self-talk contains both a speaker- and an addressee-oriented GroundP, just as typical conversations do. If two interactional roles are available, it stands to reason that the person engaged in self-talk may hold either of these two roles. This follows from the logic of dialogue in combination with the interactional spine hypothesis. Consider how. Up to this point, we have considered the interactional roles from the point of view of the person who is speaking. That is, when an individual speaks, they will be assigned the role of (holder of) Ground-Spkr (and hence the speaker role). However, when an individual listens to their interlocutor, they will interpret their interlocutor's utterance as assigning the role of Ground-Adr (and hence the addressee role) to themselves. This is schematized in (42), where ME represents the person speaking in (42a) and the person listening in (42b).

- (42) a. when I speak: [Ground-Adr [Ground-Spkr ME [p]]]
 b. when I listen: [Ground-Adr ME [Ground-Spkr [p]]]

I argue that self-talk is no different in this respect. That is, when someone engages in *you*-centered self-talk, they may identify with either of the two roles available in the structure. That is, they may identify as the speaker (and hence be assigned the speaker role) as in (43a). Alternatively, they may identify as the listener (and hence be assigned the addressee role) as in (43b).

(43) a. when I identify as the speaker in self-talk:

[_{Ground-Adr} [_{Ground-Spkr} ME [p]]]

b. when I identify as the listener in self-talk:

[_{Ground-Adr} ME [_{Ground-Spkr} [p]]]

The representation in (43) invites the question as to who would be assigned the other interactional role in self-talk. Who is the person engaged in self-talk interacting with? In other words, who is ME talking to in (43a), and who is talking to ME in (43b)? I suggest that when a person engaging in self-talk identifies as the speaker, they typically talk to an externalized image of themselves (like a picture or a mirror image), as illustrated in (44a). In contrast, when a person engaging in self-talk identifies as the listener, the role of the speaker is assumed by an internalized disembodied voice, i.e., a voice that is talking to ME, as in (44b). This voice can either be an inner critic or an inner coach, for example.¹⁵

(44) a. mirror-oriented:

[_{Ground-Adr} externalized image of me [_{Ground-Spkr} ME [p]]]

b. disembodied voice:

[_{Ground-Adr} ME [_{Ground-Spkr} disembodied inner voice [p]]]

Evidence for the distinction between these two types of *you*-centered self-talk comes from an experimental cross-linguistic study conducted by Goddard et al. (2022), who explored self-talk in English, Mandarin, and Japanese. In what follows, I focus mostly on their discussion of Japanese. The motivation for studying self-talk in Japanese (as compared to English) is rooted in the question regarding the use of socio-linguistically loaded pronouns in the context of self-talk. That is, while English has only one 2nd person pronoun, Japanese has a plethora of pronouns with complicated, sociologically determined rules of use (Kuroda 1965; see Takubo 2020 for a recent overview). According to Takubo (2020: 689f.), Japanese pronouns “cannot be freely used in conversational discourse ...[and u]nlike English, the use of second person [pro]nouns to refer to the addressee in Japanese is usually considered impolite and restricted only to individuals who are close to the speaker.” It is this property that

¹⁵ This analysis raises the question as to whether a person engaged in *you*-centered self-talk of the type in (44b) is having an auditory hallucination, i.e., whether this is a case of “hearing voices”. Note that the phenomenon of hearing voices as opposed to (covert) self-talk (inner speech) is complicated and a detailed discussion goes beyond the scope of this paper. Specifically, it has been shown that hearing voices is not only a common symptom of schizophrenia, but also found in typical populations (Fernyhough 2004). What is crucial for our purpose is that a person engaged in self-talk knows that the disembodied voice that appears to talk to them is part of their own self and is internally generated. In contrast, auditory hallucinations are due to difficulties distinguishing between internally and externally generated perceptual experiences (e.g. Johns and McGuire 1999). This is consistent with the classification of schizophrenia as a disorder of the self (Sandsten et al. 2022).

underlies the classification of Japanese as a language where pronouns are avoided for politeness (Helmbrecht 2013).

Goddard et al. (2022) observe that this socio-linguistic richness of Japanese pronouns affects self-talk. Their online survey was based on storyboards aimed to elicit spontaneous utterances as well as grammatical well-formedness judgments in self-talk. Participants were asked to either choose among various captions to go along with a particular storyboard or to provide their own captions. Storyboards targeting *you*- and *I*-centered responses were balanced across the study. Storyboards with two-person dialogues (i.e., typical conversations) served as a control. When participants had to provide their own caption, there were two notable results. First, there was a clear preference for the use of *pro*-drop in self-talk. However, these *pro*-drop utterances provide no clue as to whether they are to be classified as *I*-centered or *you*-centered self-talk. Second, when participants did produce pronouns, it was always a speaker-denoting one. Hence, these were clear instances of *I*-centered self-talk. This suggests that *you*-centered self-talk might be ill-formed in Japanese, at least with overt pronouns (though see below for a more fine-grained conclusion). This tentative conclusion is consistent with Koguma et al. (2020), who argue that in Japanese self-talk, the use of 1st person pronouns is natural, but the use of 2nd person pronouns is not. They show that this is equally true in contexts of self-encouragement, as in (45) (Koguma et al. 2020: 170), as well as self-blame, as in (46) (Koguma et al. 2020: 169).

- (45) a. ore[(w)atasi]-nara dekiru!
 I[I]-be.if can.(do it)
 ‘I can do it!’
 b. * omae[an(a)ta]-nara dekiru!
 you[you]-be.if can.(do it)
 ‘You can do it!’
- (46) a. (w)atasi[ore] nani yatten-no[daroo]?
 I what do.PROG-thing
 ‘What the heck am I doing?’
 b. *anta[omae] nani yatten-no[daroo]?
 you[you] what do.PROG-thing
 ‘What the heck are you doing?’

According to Koguma and Izutsu (2022: 20), this restriction on *you*-centered self-talk in Japanese derives from a difference in the way self-talk is conceptualized in this language. Specifically, they argue that “monologic self-reference resides in absolute solitude: the conceptualization of a speech event with no presence of addressees.”

In contrast, Goddard et al. (2022) argue that the restriction on *you*-centered self-talk in Japanese derives from the socio-linguistic richness of the 2nd person pronouns. Their use is always conditioned by the interactional context and is influenced by the identity of the addressee and the relation between the interlocutors.

However, the relevant generalization is more nuanced. Specifically, in storyboards where the character is depicted as engaging with their reflection in a mirror, a significant increase in the permissibility of *you*-centered self-talk with overt pronouns is observed (Goddard et al. 2022). Based on this finding, they conclude that the mirror provides an environment that facilitates social deixis and, hence, the use of socio-linguistically loaded 2nd person pronouns.

Now consider the Japanese facts just reviewed in light of the proposal in (44). We are led to conclude that in Japanese, *you*-centered self-talk is restricted to the mirror-oriented type, whereas the second type, where the person engaging in self-talk identifies with the listener, is unavailable in Japanese. On the present analysis, this cross-linguistic difference is grammatically conditioned rather than being dependent on a culturally conditioned conceptualization of self-talk, as proposed by Koguma and Izutsu (2022). Specifically, it reduces to a difference in the content of 2nd person pronouns between English and Japanese. Japanese 2nd person pronouns require a context that licenses social deixis. By hypothesis, an internal disembodied voice does not enter a social relation with the person engaged in self-talk. Arguably, a social relation requires an external body or at least a depiction thereof.

Independent evidence that 2nd person pronouns in English do not require that the speaker enters a social relation with the addressee comes from the fact that they do not even require the presence of an addressee, and thus, no social relation is required. This is evidenced by the fact that English 2nd person pronouns can be used as impersonal pronouns, as in (47) (Malamud 2006: 161).

(47) In those days, **you** could marry your cousin.

In contrast, in Japanese, overt addressee-denoting pronouns cannot be interpreted impersonally, as shown in (48a). In such contexts, *pro*-drop is obligatory, as in (48b) (adapted from Kitagawa and Lehrer 1990: 755).

- (48) a. Sooiu toki-ni-wa anata honnooteki-ni ugoi-te sima-u
 Such time-at-TOP you.SG instinctively moving end.up-PRS
 ‘You_{INDEXICAL}/*one react(s) instinctively at a time like that.’
 b. Sooiu toki-ni-wa *pro* honnooteki-ni ugoi-te sima-u
 Such time-at-TOP *pro* instinctively moving end.up-PRS
 ‘You_{INDEXICAL}/one react(s) instinctively at a time like that.’

According to Kitagawa and Lehrer (1990: 756), “[I]n languages like Japanese [...], the so-called (lexical) personal pronouns, especially those having to do with 1st and 2nd persons, are too closely tied to the actual speech act context. They are simply too loaded with semantic and pragmatic information.” I here suggest that the same reason is responsible for the impossibility of overt 2nd person pronouns in *you*-centered self-talk where the speaker identifies with the listener.

Additional evidence for the proposal that the cross-linguistic difference derives from the grammatical properties of pronouns, which are independently motivated, comes from Mandarin. Specifically, Mandarin has paradigmatic pronouns of the type found in English, but the pronominal paradigm also contains a formal addressee-denoting pronoun (*nín*). This formal pronoun is similar to Japanese pronouns in that its use is socio-linguistically constrained. However, it also differs from Japanese pronouns as it is not part of a rich inventory of socio-linguistically loaded forms. It is a unique formal pronoun akin to those found in languages with a *tu/vous* distinction. Significantly, Goddard et al. found that the Mandarin formal pronoun, like Japanese addressee-denoting pronouns, is more felicitous in the context of mirror-oriented self-talk than it is in the context of engaging with a disembodied voice. In contrast, the unmarked addressee-denoting pronoun does not show these differences.

The same generalizations appear to hold in German and French: the formal pronoun is restricted to mirror-oriented *you*-centered self-talk, while the unmarked pronoun can be used in both types of self-talk. This is illustrated below for German.

- (49) Sie haben das gut gemacht.
 you_{formal} have that well done
 ‘You did well.’
 ✓ mirror-oriented self-talk
 ✗ self-talk with disembodied voice
- (50) Du hast das gut gemacht.
 you have that well done
 ‘You did well.’
 ✓ mirror-oriented self-talk
 ✓ self-talk with disembodied voice

Taken together, these facts, summarized in Table 4, suggest that restrictions on different modes of self-talk are grammatically rather than culturally conditioned. The initial observation, due to Koguma et al. (2020), appears to be a difference between English and Japanese because these two languages have different types of pronouns: Japanese pronouns contain social deixis, but English pronouns do not. If we extend the cross-linguistic exploration to include

Table 4: Constraints on modes of self-talk.

	<i>I</i> -centered self-talk	<i>you</i> -centered self-talk	
		Mirror-oriented ME = speaker	Disembodied voice ME = listener
English	✓	✓	✓
Japanese	✓	✓	✗
Mandarin unmarked pronoun	✓	✓	✓
Mandarin formal pronoun	✓	✓	✗
German unmarked pronoun	✓	✓	✓
German formal pronoun	✓	✓	✗

languages that have pronouns with and without social deixis, we observe that it is indeed the type of pronoun that restricts the availability of *you*-centered self-talk. Thus, it cannot be viewed as a culturally determined language-wide difference.

4.3 A grammatical difference between self-talk and typical conversations

At the core of the present analysis lies the claim that self-talk comes in two guises: *I*-centered self-talk is essentially a form of thinking out loud, while *you*-centered self-talk corresponds to having a conversation with oneself. This difference, I argued, is structurally conditioned in that only *you*-centered self-talk contains a syntactic position for the addressee (SpecGround_{Adr}P). If *you*-centered self-talk is indeed a conversation with oneself, the question arises as to whether it differs from a typical conversation with another individual. This is the question I address in this subsection. I show that there is at least one difference involving the use of intonational tunes, and I argue that this difference is structurally conditioned (see Ritter and Wiltschko 2021).

First, consider the intonational tunes associated with *wh*-questions in a typical conversation. As shown in (51), in a typical conversation, *wh*-questions can be realized with either falling (↘) or rising (↗) intonation (Bartels 1999; Bolinger 1989).

- (51) Alaka to Theo:
- a. What are you doing ↗
 - b. What are you doing ↘

Crucially, *you*-centered self-talk differs in this respect, as shown in (52). Only falling intonation is possible when talking to oneself (Ritter and Wiltschko 2021; see also Krifka 2023b).

- (52) Alaka to himself:
 a. * What are you doing ↗
 b. What are you doing ↘

Following Wiltschko and Heim (2016), I assume that a rising intonational tune signals a call on the addressee to respond. As such, it occupies the head of Resp(onse)P, the highest category in the interactional structure.¹⁶

Falling intonation, in contrast, is the default and, hence, is not interpreted as a meaningful intonational tune. This is because it arises naturally due to the fact that during an utterance, pitch declines automatically with the decrease in subglottal air pressure (Cohen et al. 1982). As such, unlike rising intonation, falling intonation does not have a syntactic representation (Wiltschko 2024a).

Given this analysis, the impossibility of rising intonation in *you*-centered self-talk indicates that RespP is available in a typical conversation, as in (53a), whereas it is not in *you*-centered self-talk, as in (53b) (Ritter and Wiltschko 2021). Hence, there is no position available to host rising intonation, as illustrated in (53b) by means of the asterisk.

- (53) a. [RespP ↗ [Ground-Adr [Ground-Spkr [p]]]] typical conversation
 b. * [Ground-Adr [Ground-Spkr [p]]] *you*-centered self-talk

This analysis implies that in self-talk, the grammatical regulation of turn-taking is not available. While in typical conversations, the addressee is an active participant from whom the speaker can request a response, in *you*-centered self-talk, they are not. A person engaged in self-talk does not call on themselves to respond. This analysis is consistent with Holmberg's (2010: 57) observation that self-talk is always a "one-way communication."

The logic behind Ritter and Wiltschko's analysis regarding the lack of RespP in self-talk holds for initiating moves only, where the speaker may signal a request for a response by their interlocutor. This is implemented by assuming that the Response-set is other-oriented, as in (30b), repeated below for convenience. The situation is, however, different in reaction moves, where the response set is self-oriented (30a).

¹⁶ As is well-known, there are several different types of rising tunes (Goodhue 2024; Heim 2019). The discussion here refers to the rise associated with questionhood.

- (30) a. Reaction move:
 [_{RespP} Resp-set_{Self} [_{GroundAdrP} Ground-Adr [_{GroundSpkrP} Ground-Spkr [p]]]]
 b. Initiating move:
 [_{RespP} Resp-set_{Other} [_{GroundAdrP} Ground-Adr [_{GroundSpkrP} Ground-Spkr [p]]]]

Crucially, reaction moves can but need not be reactions to previous conversational moves. Rather, speakers can react to non-linguistic events as well, and if they do, RespP is still available (Wiltschko 2021, 2024b). In the present context, this raises the question as to whether RespP is available in self-talk when it serves to mark a reaction move. In other words, is the structural deficiency of *you*-centered self-talk purely grammatically conditioned, or does it reflect the logic of initiation versus reaction moves? If it is purely grammatically conditioned, we would expect a categorical absence of RespP, no matter whether it marks initiation or reaction. If, instead, it is pragmatically conditioned, we would not expect a categorical ban on RespP, but instead, RespP should be available when it serves to mark the move as a reaction move. In what follows, I present evidence suggesting that RespP is available when a person engaged in self-talk reacts to a non-linguistic event.

Upper Austrian German has a discourse marker (*ma*), which serves to mark a reaction move (Wiltschko 2024b). It can be used to mark a reaction to a non-linguistic event, as in (54) (Wiltschko 2024b: 183).

- (54) Context: Xaver sees his friend drawing a beautiful picture. He is surprised that his friend can draw.
 Xaver: **Ma** was moch-st'n du do schens?
 MA what make-2SG=PRT 2SG there beautiful
 ‘What beautiful thing are you making?’ (indicating surprise)

Wiltschko (2024b) analyses *ma* as a pro-form for a speaker-oriented GroundP, which occupies the specifier position of a self-oriented (reacting) RespP, as schematized in (55).

- (55) [_{RespP} [_{GroundSpkr} *ma*] Resp_{SELF} [...]]

According to this analysis, with the use of *ma*, a speaker marks a reaction to their own epistemic state. This, in turn, captures the fact that *ma* may indicate surprise. However, *ma* is not a discourse marker dedicated to surprise. Rather, *ma* can also be used in a context where the speaker suddenly remembers something, as in (56) (Wiltschko 2024b: 193).

- (56) Context: Reingard suddenly remembers that she needs to return a book to Mariana. So Reingard tells Mariana:
 Ma do foit ma grod ei, I woit da dei buach zruckgem
 MA there falls 1SG.DAT just in, I wanted 2SG.DAT 2SG.POSS book return
 ‘I just remembered: I wanted to return your book.’

In this context, no surprise is involved as the propositional content is something that Reingard previously knew but has temporarily forgotten. What the two contexts in (54) and (56) have in common is that the speaker marks a reaction to their own epistemic state. Thus, *ma* constitutes an ideal test case for whether self-talk allows for the presence of RespP in reaction moves. Significantly, *ma* can be felicitously used in self-talk, no matter whether it is *I*-centered, as in (57a), or *you*-centered, as in (57b).

(57) Xavier to himself upon realizing that he forgot his phone at home.

- a. **Ma** I bin a gonza Depp.
MA I am a whole idiot
 ‘Geez, I’m a real idiot.’
- b. **Ma** Xavier du bist a gonza Depp.
MA Xavier you are a whole idiot
 ‘Geez, Xavier, you are a real idiot.’

For completeness, note that English too, has discourse markers used in similar contexts. For example, Schourup (1982: 14) analyses *oh* as marking the presence of unspoken thought (see also Aijmer 1987). He provides the minimal pair in (58) (Schourup 1982: 15), which illustrates the contribution of *oh* within a typical conversation.

- (58) a. I didn’t make the phone call you asked me to.
- b. **Oh!** I didn’t make the phone call you asked me to.

According to Schourup (1982: 15), *oh* in (58b) indicates that the thought expressed in the following sentence just entered the speaker’s mind. Thus, with the use of *oh* a speaker may implicate that their failure to make the call was due to forgetfulness and not malevolent intent.

This description suggests that *oh* might be amenable to an analysis akin to Wiltschko’s (2024b) analysis of Upper Austrian *ma*. Specifically, *oh* is used to mark the utterance as a reaction and hence it may be analyzed as occupying SpecResp. If so, the examples in (59) provide evidence from English too, that RespP is licensed in self-talk, at least in the context of reaction moves.

- (59) Alaka to himself upon realizing that he forgot to call Thea:
- a. **Oh**, I’m such an idiot. I forgot to call Thea.
- b. **Oh**, Alaka, you are such an idiot. You forgot to call Thea.

In sum, we have now seen that *you*-centered self-talk differs from typical conversations only in that the need to regulate turn-taking is obviated. Consequently, the means to do that are not available, as evidenced by the impossibility of rising intonation in questions. This suggests that initiation moves are structurally deficient in self-talk with RespP missing. However, we have also seen that there is no categorical restriction on the projection of RespP, as evidenced by the fact that reaction moves

can be introduced by units of language associated with RespP. This supports the view that self-talk may, in fact, be fundamentally interactional, just as regular conversations are, even though there is no other interlocutor to interact with.

4.4 Interim conclusion: an interaction-based typology of self-talk

In this section, we have seen empirical evidence for a three way-distinction in self-talk: *I*-centered self-talk differs from *you*-centered self-talk in that it does not license any linguistic phenomena that require the grammatical representation of an addressee (vocatives, imperatives, and addressee-oriented discourse markers). On the other hand, *I*-centered but not *you*-centered self-talk allows for the use of verbs of cognition. In addition, we have seen evidence for two types of *you*-centered self-talk. Specifically, the person engaged in self-talk can take on the role of the listener, with the speaker being construed as a disembodied (inner) voice. Alternatively, the person engaged in self-talk can take on the role of the speaker with an externalized picture or mirror image serving as the addressee. These two types of *you*-centered self-talk are empirically distinguished in the availability of social deixis (e.g., formal addressee-denoting pronouns): a disembodied voice is not a social being, and hence social deixis is not allowed when the person engaged in self-talk takes on the role of a listener. Finally, self-talk differs from typical conversations in that there is no need for the regulation of turn-taking, and hence, rising intonation is not available, though reactions can be marked as such.

These empirical differences (summarized in Table 5) receive a straightforward analysis within an interaction-based approach toward the topmost structure of an

Table 5: Modes of talking: empirical differences.

	<i>I</i> -centered self-talk	<i>you</i> -centered self-talk		Typical conversation
		ME = listener	ME = speaker	
Vocatives	✗	✓	✓	✓
Imperatives	✗	✓	✓	✓
Addressee-oriented discourse markers	✗	✓	✓	✓
Verbs of cognition	✓	✗	✗	✗
Social deixis	✗	✗	✓	✓
Rising intonation	✗	✗	✗	✓
Markers of reaction	✓	✓	✓	✓

Table 6: Modes of talking: structural differences.

	<i>I</i> -centered self-talk	<i>you</i> -centered self-talk	Typical conversation
Ground-Spkr	✓	✓	✓
Ground-Adr	✗	✓	✓
Resp SELF	✓	✓	✓
Resp OTHER	✗	✗	✓

utterance, as summarized in Table 6. The difference between *I*-centered and *you*-centered self-talk resides in the presence versus absence of the addressee-oriented grounding phrase. The difference between typical conversations and self-talk resides in the presence versus absence of the OTHER-oriented response phrase, the layer of structure that hosts the request for a response in initiating moves.

On this approach, the differences in modes of talking are structurally conditioned. The difference between the two types of *you*-centered self-talk derives from the assumption that the person engaged in self-talk may identify with either the listener or the speaker and hence be assigned the role of holder of Ground-Spkr or Ground-Adr, which are both available in *you*-centered self-talk. In this way, *you*-centered self-talk is no different from typical conversations, in which an individual too, may identify as the speaker or as the listener.

Significantly, what the empirical properties of self-talk reveal and what the structural analysis captures is that human language does not seem to have a dedicated means for self-talk. According to the analysis developed here, the different modes of self-talk reflect structural differences. But crucially, there is no dedicated structure for self-talk.¹⁷ If there were a dedicated structure for self-talk, we might expect it to take on the form of a reflexive construction of sorts. That is, languages have means to mark that event roles are assigned to one individual only and hence that the referent holding the *agent* role is identical to the referent holding the *patient* role. This type of reflexive marking in the event domain can be realized on the predicate or the internal argument in the form of a reflexive pronoun (Reinhart and Reuland 1993). To the best of my knowledge, there are no equivalent reflexive

¹⁷ Even in addressee-oriented regular conversations the full interactional structure is not always projected. Rather, Wiltschko (2021) argues that the presence of interactional structure is regulated by assumptions about the normal course of a conversation. For example, under normal circumstances, when a speaker answers a question, it need not be marked as a response. It simply functions as a response by virtue of its occurrence in the conversation. However, a response needs to be marked as such if it is an unexpected response (i.e., if it does not directly provide an answer, for example). See Wiltschko (2021) for full discussion.

markers available for interactional roles. In other words, I know of no markers that would indicate that the individual holding the speaker role is identical to the individual holding the addressee role. On this view, then, self-talk itself is not special. Rather, the differences in grammatical representation are independently available. The structure associated with *I*-centered self-talk simply corresponds to a structure that represents a thought an individual may hold and express. Thus, it can be conceptualized as thinking out loud. In contrast, the structure associated with *you*-centered self-talk corresponds to a structure that represents a thought packaged for conversational interaction (i.e., embedded in a structure that makes the utterance sensitive to the presence of an addressee). In this way, *you*-centered self-talk can be conceptualized as having a conversation with oneself. And like in a typical conversation, the person engaged in this conversation may identify as the speaker or the listener.

Suggestive evidence for this typology of self-talk and for the assumption that the different modes of talking are associated with different grammatical and, thus, mental representations comes from the fact that the same distinctions are observed in covert self-talk (i.e., the phenomenon often referred to as inner speech, see Section 2.1).

Specifically, it is known that covert self-talk is not a unified phenomenon but instead that it is subject to substantial variability (Hurlburt et al. 2013). On the one hand, there is the often-discussed difference between condensed and expanded inner speech (Fernihough 2004), which has no direct correlate in overt self-talk as the latter cannot be condensed. However, the differences of the type discussed here for overt self-talk have also been observed for covert self-talk. In the case of covert self-talk, the evidence does not come from linguistic considerations, however, but instead is based on neuroimaging data.

First, based on a neuroimaging study, Alderson-Day et al. (2016) establish a difference between monologic and dialogic covert self-talk. They show that in dialogical versions of covert self-talk, a wider network is involved than the classical regions associated with language (production and comprehension). Specifically, dialogic covert self-talk appears to involve regions responsible for Theory of Mind and social cognition. This appears to correlate with the distinction between *I*-centered and *you*-centered self-talk we have established for overt self-talk based on linguistic evidence.

In addition, Hurlburt et al. (2016) point out that in the literature, the term inner speech may refer to two phenomenologically and psychologically distinct phenomena, which they refer to as *inner speaking* and *inner hearing*, respectively. From the present point of view, this seems to correspond to the difference between the person engaged in overt self-talk identifying with the speaker or with the listener. This is corroborated by neuroimaging results that show evidence for two types of inner

speech, one where production-related regions are activated, whereas another where perception-based regions are activated (Pratts et al. 2023). According to Kompa (2024: 647), “[t]here is inner speaking and signing, which is accompanied by a sense of agency, and inner hearing or auditory imagery [...], where one experiences oneself as being passive.” Fernyhough (2009) suggests that the latter type of self-talk is akin to a dialogue between the person engaged in self-talk, who identifies as a listener, and an inner interlocutor who is conceptualized as an ‘open slot’ fillable by any imaginary interlocutor.

I conclude that the typology of self-talk we have established here based on linguistic evidence has psychological correlates. These psychological correlates point towards the significance of viewing (some forms of) self-talk as intrinsically dialogical. I thus submit that an interaction-based analysis of the type developed here is well suited to capture not only the linguistic profiles of the different types of self-talk but also their psychological characteristics.

5 Alternative analyses

We now turn to alternative analyses of self-talk within other frameworks that incorporate syntactic structure at the very top, as discussed in Section 3. I start with a discussion of Krifka’s (2023b) analysis of self-talk framed within a commitment-based approach (Section 5.1), and I point out some empirical and conceptual problems it faces. I then discuss how the properties of self-talk might be analyzed within a neo-performative approach (Section 5.2).

5.1 A commitment-based analysis of self-talk

To the best of my knowledge, Krifka (2023b) presents the only other explicit linguistic analysis of self-talk within a framework that assumes that some pragmatic properties are syntactically encoded at the very top of the tree. Drawing on the insights of Ritter and Wiltschko (2021) and using data from Schnitzler’s *Leutnant Gustl*, he also analyses the difference between *I*-centered and *you*-centered self-talk in terms of structural deficiency. At the core of his proposal is the claim that *I*-centered self-talk lacks CommitP, i.e., the phrase dedicated to encoding a speaker’s (public) commitment to the propositional content. Rather, *I*-centered self-talk consists of only JudgeP (encoding the speaker’s private perspective) and ActP (encoding the illocutionary force). This contrasts with *you*-centered self-talk, where CommitP is present, just as in typical conversations. This is schematized in (60) (adapted from Krifka 2023b).

- (60) a. *I*-centered self-talk: [ActP Act [JudgeP Judge [TP p]]]
 b. *you*-centered self-talk: [ActP Act [CommitP Commit [JudgeP Judge [TP p]]]]

Note that the core insight behind the interaction-based analysis had to do with the absence of the structural position responsible for introducing the addressee role (Ground_{Adr}). Crucially, Krifka's commitment-based structure does not incorporate such an addressee-oriented phrase. Rather, in this framework, the corresponding structural positions distinguish between a private perspective and a public commitment, where the public nature of commitment arguably requires an addressee (otherwise, the commitment would not be public). On this view, then, the difference between *I*- and *you*-centered self-talk amounts to a difference between expressing one's private thought (i.e., thinking out loud) and publicly committing to it (i.e., having a conversation with oneself). While this captures the core insight of the interaction-based analysis, in the absence of reference to an addressee in CommitP, the ungrammaticality of vocatives, imperatives, and other addressee-oriented units of language does not straightforwardly follow. This is not to say that the analysis could not be amended to capture these empirical differences, but it does not follow from the structural analysis alone. To be sure, the additional contextual variables that are incorporated in Krifka's analysis will not help in this respect. Recall that Krifka's commitment-based analysis is amended through contextual variables that introduce the relevant roles (speaker, addressee, judge, and committer) and stipulations of identity. Typically, the judge is identified as the committer. However, the identity of the committer differs depending on the illocutionary force. In assertions, the committer is identified as the speaker, while in questions, it is identified as the addressee. The relevant representations for assertions and questions in typical conversations are repeated below for convenience (Krifka 2023b: 350f.).

- (28) Assertion:
 [ActP • c:=s[CommitP c commits to j:=c[JudgeP j judges [TP ...]^{j,s,a}]^{s,a}]^{s,a}]
- (29) Question:
 [ActP ? c:=a[CommitP c commits to j:=c[JudgeP j judges [TP ...]^{j,s,a}]^{s,a}]^{s,a}]

To capture the essence of self-talk, Krifka (2023b) amends the contextual variables. Specifically, *I*-centered self-talk is characterized by the absence of an addressee (hence, the contextual variable for the addressee is marked with strike-through in (61a)). In addition, Krifka proposes that the participants in *you*-centered self-talk can be identified as an *inner self* ('Inneres Selbst' IS) and an *outer self* ('Äußeres Selbst' ÄS). The inner self can have thoughts and feelings (expressed in *I*-centered self-talk), and the outer self can reflect upon these thoughts and is thus identified as the locus of

identify either with the speaker or with the listener. This option is not available in the commitment-based analysis. If we were to speculate that, in this case, the person engaged in self-talk could identify with either the inner or the outer self, this would not explain why there should be a difference in the licensing of social deixis. Rather, we would expect that social deixis is ruled out in both types of *you*-centered self-talk since neither the inner nor the outer self can be viewed as social beings relative to each other.

A final empirical challenge the commitment-based analysis faces is that it fails to account for the difference in the availability of rising intonation. In the interaction-based analysis, this difference can be modelled through the availability of the topmost structure (RespP) in initiating moves. The commitment-based approach towards the very top of the tree has no equivalent structure. Thus, this difference must be accommodated differently.

Lastly, there is also a conceptual disadvantage the commitment-based analysis faces. It relies on the postulation of contextual variables that only seem to play a role in self-talk. That is, the variables for speaker and addressee available in typical conversations are replaced by inner and outer self, which are, by hypothesis, available in self-talk only. This amounts to saying that self-talk comes with its own grammatical ingredients. In contrast, the interaction-based analysis did not postulate any dedicated means for self-talk. I submit that this is a conceptual advantage of the interaction-based analysis over the commitment-based one.

5.2 A neo-performative analysis of self-talk

I now turn to a discussion of the neo-performative framework and its potential to analyze the properties of self-talk. To the best of my knowledge, there is no explicit analysis of self-talk available to draw upon. Hence, the discussion is restricted to what such an analysis might look like (subsection 5.2.1). There is, however, a neo-performative analysis available for self-directed questions. Given the apparent affinity between self-directed questions and self-talk, I shall include a discussion of this phenomenon as well (subsection 5.2.2).

5.2.1 What a neo-performative analysis of self-talk might look like

The first thing to note is that a neo-performative analysis has the ingredients necessary to distinguish between *I*-centered and *you*-centered self-talk. The neo-performative structure, repeated below for convenience, contains a speaker and an addressee role.

- (25) $[_{saP} \text{ Spkr } [sa] [_{saP} p [_{sa} [sa] \text{Adr}]]]$

Thus, on a neo-performative approach, we might analyze the difference between *I*- and *you*-centered self-talk as resulting from the absence or presence of the addressee role, as shown in (62).

- (62) a. *I*-centered self-talk: $[_{saP} \text{ Spkr } [sa] [_{saP} p [_{sa} [sa]]]]$
 b. *you*-centered self-talk: $[_{saP} \text{ Spkr } [sa] [_{saP} p [_{sa} [sa] \text{Adr}]]]$

This type of analysis would keep in line with the assumption that *I*-centered self-talk is structurally deficient compared to *you*-centered self-talk, though the structural difference is qualitatively different from the one assumed in the interaction-based analysis. Rather than lacking a functional projection, what is missing is an object. That is, in the neo-performative analysis of Speas and Tenny (2003), the speaker functions as the subject and the addressee as the indirect object of the speech act structure, with the utterance content (p) functioning as a direct object. Thus, the difference between *I*- and *you*-centered self-talk would be akin to the difference between a transitive and a ditransitive argument-structure. However, while the insight that *I*-centered self-talk is characterized by the absence of an addressee can be maintained, it is less clear if and how the empirical correlates we have observed would follow on this analysis. To capture the difference between self-talk and typical conversations (i.e., the availability of rising intonation in questions), the neo-performative analysis would have to be amended. This is because, unlike the interaction-based framework, there is no layer of structure dedicated to regulating turn-taking. This is not to say that an analysis of these empirical differences in the modes of talking is impossible on a neo-performative-based account, but it is arguably less straightforward than it is under an interaction-based account.

A more serious problem the neo-performative approach faces has to do with one of its core assumptions, namely that the structure at the very top regulates speech acts. Neo-performative approaches typically assume that this structure is, in fact, the locus of illocutionary force. As we have seen in Section 3, Speas and Tenny (2003) assume that the structure in (25) corresponds to a declarative clause and is interpreted as an assertion (26a), whereas interrogatives are derived by moving the addressee role to a position where it takes scope over p, as in (26b). This analysis is meant to reflect the fact that the epistemic authority is now with the addressee (adapted from Speas and Tenny 2003: 320).

- (26) a. Declarative: $[_{saP} \text{ Spkr } [sa] [_{saP} p [_{sa} [sa] \text{Adr}]]]$
 b. Interrogative: $[_{saP} \text{ Spkr } [sa] [_{saP} \text{Adr } [p [_{sa} [sa] \text{Adr}]]]]$

Everything else being equal, we would expect that the absence of the addressee role in *I*-centered self-talk would make it impossible to derive questions or even

to distinguish between assertions and questions. This is, however, not the case. Specifically, we have already seen examples of declaratives used in *I*-centered self-talk in Section 2.3. An example is repeated below for convenience.

- (63) Alaka to himself:
- a. I am an idiot.
 - b. I will go to the gym today, no matter what.

In addition, *I*-centered self-talk also allows for questions (both polar questions and *wh*-questions), as shown in (64).

- (64) Alaka to himself:
- a. Am I crazy now?
 - b. What am I doing here?

Recall that based on the distribution of vocatives, imperatives, and addressee-oriented discourse markers, we know that *I*-centered self-talk is indeed defined by the absence of a grammatically represented addressee role. Hence, given the availability of interrogatives in *I*-centered self-talk, we can conclude that their well-formedness does not depend on the presence of an addressee role. Instead, questionhood must be intrinsic to the propositional structure. This is consistent with the assumption that questioning is a type of propositional attitude – in other words, attitudes can have questions as their content (Friedman 2013).¹⁸

This conclusion, however, undermines one of the core motivations for Speas and Tenny's neo-performative approach, which is to derive the fact that the number of clause-types is universally restricted. Again, this is not to say that a neo-performative-based approach cannot be adjusted to account for the typology of modes of talking we have observed here. But the fact remains that more needs to be said to do so. And in fact, some neo-performative approaches deal with this issue. That is, while there is to date no explicit analysis of self-talk available within a neo-performative approach, there is a closely related phenomenon for which there is and to which we now turn.

5.2.2 A neo-performative analysis of conjectural questions

Across different languages, there are questions which appear to be dedicated to being self-addressed (see Truckenbrodt 2006 for German, Littell et al. 2010 for Amerindian languages, Oguro 2017 for Japanese, and Eckardt and Disselkamp 2019

¹⁸ We are facing terminological problems here. By questioning I here refer to the propositional attitude. It is to be distinguished from the speech act notion of a question, which requires an addressee.

for Korean). Such questions have variously been referred to as *self-addressed*, *hearerless*, *deliberative*, or *conjectural*. Here, I adopt the term *conjectural*. Conjectural questions are typically associated with a dedicated syntactic form. For example, in German, they are characterized by verb-finality and are often marked with the discourse marker *wohl*, as in (65a) (adapted from Krifka 2013:15). In contrast, typical (true) questions are characterized by verb-second and are often marked with a different discourse marker (*denn*), as in (65b).

- (65) a. Wo der Schlüssel wohl wieder sein mag?
 where DET key WOHL again be may
 ‘[I wonder] where the key might be.’
 b. Wo ist denn wieder der Schlüssel?
 where is DENN again DET key
 ‘Where is the key?’

In Korean, conjectural questions are characterized by a special sentence-final particle (*na*), as in (66a), which differs from the particle used in true questions (*ni*), as in (66b) (Eckardt and Disselkamp 2019: 383).

- (66) a. Mary-ka o-ass na
 Mary-NOM come-PAST SAQ
 ‘Has Mary come, I wonder.’
 b. Mary-ka o-ass ni?
 Mary-NOM come-PAST TRUEQ
 ‘Has Mary come?’

Finally, in Japanese, conjectural questions are characterized by the absence of the politeness marker *masu*, as in (67a) (Oguro 2017: 192), which is otherwise obligatory in matrix questions, as in (67b) (Miyagawa 2012:15).

- (67) a. Dare-ga ku-ru ka?
 who-NOM come-PRS Q
 ‘(I wonder/I’m not certain) Who will come.’
 b. Dare-ga ki-masu ka?
 who-NOM come-POLITE Q
 ‘Who will come?’

Conjectural questions are relevant in the present context because there is a neo-performative analysis available, which in turn might be applicable to self-talk more generally. That is, given that conjectural questions are sometimes described as being used in the absence of an interlocutor, i.e., in a monologue (Jang and Kim 1998; Jang 1999), they appear to be classifiable as a form of *I*-centered self-talk in the sense discussed here.

This conclusion is corroborated by the observation that conjectural questions are incompatible with *you*-centered self-talk (Eckardt and Disselkamp 2019; Krifka 2023b), as shown for German in (68) (Krifka 2023b: 17) and for Korean in (69) (Eckardt and Disselkamp 2019: 392).

- (68) a. Wo ich wohl wieder den Schlüssel hingelegt habe.
 where I WOHL again DET key put have.1SG
 ‘(I wonder) where I put the key again.’
 b. ??Wo du wohl wieder den Schlüssel hingelegt ha-st
 where you WOHL again DET key put have-2SG
 ‘(I wonder) where you put the key again.’
- (69) a. Ney yelsay-ka eti(-ye) iss-ni?
 your key-NOM where(-LOC) exist-TRUEQ
 ‘Where is your key?’
 b. *Ney yelsay-ka eti(-ye) iss-na?
 your key-NOM where(-LOC) exist-SAQ

Significantly, Oguro (2017) provides a neo-performative structural analysis for conjectural questions; he proposes that they are characterized by the absence of an addressee role in the speech act structure. Thus, Oguro’s analysis is of the type I have hypothetically proposed for *I*-centered self-talk (i.e., (62a)). What the absence of the addressee does, according to Oguro (2017), is that it suppresses the information-seeking aspect of questions, deriving their interpretation as conjectural.

If this were the case, however, we would expect the same interpretation in interrogatives that are formed as conjectural questions and those that are formulated as true questions when used in *I*-centered self-talk. This is because the latter, too, is characterized by the absence of a grammatically represented addressee, as I have shown. This is, however, not the case. According to Truckenbrodt (2004), there is a crucial difference between true questions (even when uttered in self-talk) and conjectural questions. Roughly, true questions are used when the speaker thinks an answer might be available, whereas conjectural questions are used when this is not the case. This is illustrated for typical conversations with the minimal pairs in (70) and (71). When the speaker has reason to think that the addressee has no answer to their question, conjectural questions, but not true questions, are well-formed, as in (70). In contrast, when the speaker has reason to believe that the addressee might have an answer, then true questions, but not conjectural questions, are well-formed, as in (71).

- (70) Alaka and Theo are in a 2nd hand store. They encounter a strange contraption. Neither of them knows what it is. So, Alaka says to Theo:
 a. Was man damit wohl macht.
 what IMPERS there-with WOHL make-3SG
 ‘What does one do with this, I wonder.’

- b. # Was macht man (denn) damit?¹⁹
 what make-3SG IMPERS DENN there-with
 ‘What does one do with this?’
- (71) Alaka is in a 2nd hand store. He encounters a strange contraption. He doesn’t know what it is. But he assumes that the store owner will know. So, Alaka says to the store owner:
- a. # Was man damit wohl macht.
 what IMPERS there-with WOHL make-3SG
 ‘What does one do with this, I wonder.’
- b. Was macht man (denn) damit?
 what make-3SG IMPERS DENN there-with
 ‘What does one do with this?’

Note that the data in (70) already falsifies the claim that the defining feature of conjectural questions is the absence of an addressee (see Eckardt 2020).

Crucially, for our purpose, we observe a similar contrast in *I*-centered self-talk, casting doubt on the hypothesis that the interpretation of conjectural questions derives from the absence of an addressee. To see this, consider the contrast between (72) and (73). In (72) the person engaging in self-talk is faced with a decision that they have control over and hence an answer is possible. In this context, the conjectural question is infelicitous, while the regular question is well-formed. In contrast, in (73) the person engaging in self-talk is wondering about something they do not have full control over and in this context, both conjectural and regular questions are well-formed.

- (72) Alaka is browsing in a new concept store, and he loves many items. He definitely wants to buy something but is not sure what. So, he says to himself:
- a. # Hmmm... Was ich wohl kaufen soll.
 Hmmm... what I WOHL buy shall
 ‘Hm... what should I buy, I wonder.’
- b. Hmmm... Was soll ich kaufen?
 Hmmm... what shall I buy
 ‘Hm... what shall I buy?’
- (73) Alaka had a streak of unusual things happening to him for the last few days. He wakes up in an anticipatory mood, wondering what the day will bring. So, he says to himself:

¹⁹ There is a well-formed version of this utterance in self-talk, which involves the incredulity contour and stress on DAMit. I abstract away from such forms as they do not classify as conjectural questions.

- a. Was ich wohl heute wieder erleben werde.
What I WOHL today again live will
'What will happen to me again today, I wonder.'
- b. Was werde ich heute wieder erleben?
What will I today again live
'What will happen to me again today?'

The contrast between (72) and (73) suggests that conjectural questions are ill-formed when their answer pertains to something that is in the speaker's epistemic control, at least in principle, while no such constraint is associated with regular questions. This leaves us with the conclusion that the interpretation of conjectural questions cannot be characterized by the absence of an addressee (contra Jang and Kim 1998; Oguro 2017) but rather by the fact that there cannot be a straightforward answer. More precisely, I assume, following Eckardt (2020), that conjectural questions ask for answers that are "defeasibly inferred."²⁰ This derives the fact that conjectural questions are restricted to contexts where there is no direct evidence (and hence no certainty) that could be used to answer the question posed, i.e., any potential answer can only be inferred. Significantly, this aspect of meaning appears to be dependent on a feature in C, which prevents verb movement to C, and which thus results in the fact that conjectural questions are realized with the verb in final position.²¹ Thus, the difference between conjectural and true questions resides within the propositional structure and hence does not correlate with structural deficiency at the top of the tree.

We can now conclude that Oguro's (2017) analysis of conjectural questions cannot be on the right track, and as a consequence, it cannot save the problems the neo-performative analysis faces in light of self-talk. Specifically, questioning does not depend on the presence of an addressee.

6 Summary, conclusions, and future research

The goal of this paper was to explore the linguistics of self-talk and thus introduce this phenomenon as a fruitful empirical domain for linguists, as it has the potential to shed light on several theoretical questions. We started with the observation, first

²⁰ More precisely, Eckardt (2020: 35) proposes that "conjectural questions ask for answers that are defeasibly inferred from pooled knowledge of speaker and addressee". Given the availability of conjectural questions in *I*-centered self-talk, we have to conclude that the reference to speaker and addressee is not grammatically significant.

²¹ Absence of verb movement to C is not a necessary condition of the conjectural interpretation. The addition of *wohl* alone seems to suffice (cf. Eckardt 2020: 1; Zimmermann 2013: 1).

introduced into linguistic scholarship by Holmberg (2010), that a person engaged in self-talk may refer to themselves with *I* or with *you*. These two modes of self-talk are associated with distinct linguistic profiles. Specifically, *I*-centered self-talk differs from *you*-centered self-talk in that it is not compatible with phenomena that require the grammatical representation of an addressee role (vocatives, imperatives, and addressee-oriented discourse markers). Conversely, only *I*-centered but not *you*-centered self-talk allows for the use of verbs of cognition. I argued that this provides evidence for the grammatical representation of the addressee role in *you*-centered self-talk: the addressee role is always construed as occupied by an *OTHER*, whose mind is not accessible to the speaker. This is so even when real-world knowledge suggests otherwise (i.e., the person engaged in self-talk will always have access to their own mind). Moreover, based on evidence from the use of 2nd person pronouns that are socio-linguistically loaded, we have distinguished between two modes of *you*-centered self-talk: one is oriented towards an external representation of the self (e.g., a mirror image), which is treated as the addressee and the second is conceptualized as an internal disembodied voice talking to the self. Only the former is compatible with social deixis, arguably because social deixis requires a social body, which, by hypothesis, can be approximated by an image, but an inner voice cannot. Finally, we have seen that turn-taking need not and cannot be regulated in self-talk. Hence, rising intonation, a marker of calling on the addressee to respond in English typical conversations, is ill-formed in self-talk. Nevertheless, explicit markers of reactions (such as utterance-initial *oh*) are well-formed in all modes of talking, including self-talk. This is because marking a reaction is not restricted to the reaction to a prior turn by an interlocutor but can also be used in reaction to a non-linguistic event. These differences among the modes of talking are summarized in Table 5, repeated below for convenience.

Table 5 Modes of talking: empirical differences.

	<i>I</i> -centered self-talk	<i>you</i> -centered self-talk		Typical conversation
		ME = listener	ME = speaker	
Vocatives	✗	✓	✓	✓
Imperatives	✗	✓	✓	✓
Addressee oriented discourse markers	✗	✓	✓	✓
Verbs of cognition	✓	✗	✗	✗
Social deixis	✗	✗	✓	✓
Rising intonation	✗	✗	✗	✓
Markers of reaction	✓	✓	✓	✓

The linguistic evidence thus suggests that *you*-centered self-talk may be conceptualized as having a “conversation with oneself” (i.e., an inner dialogue), where the person engaged in self-talk may identify with either the speaker or the listener. In contrast, the obligatory absence of the addressee role in *I*-centered self-talk suggests that it may be conceptualized as “thinking out loud” (i.e., an inner monologue).

The linguistic properties of self-talk thus provide evidence for the grammatical representation of the addressee role, an assumption that has gained traction in linguistic theory over the past two decades. As I have shown, the distinction between *I*- and *you*-centered self-talk would be difficult to reconcile on a purely pragmatic account, but it receives a straightforward analysis on the assumption that the addressee role is introduced in the structural representation of an utterance. Moreover, I have shown that an interaction-based account allows for an analysis in terms of structural deficiency, such that *I*-centered self-talk is characterized by the absence of the addressee-oriented structural position. In contrast, other frameworks that assume the syntactic representation of speech act-based notions, such as Krifka’s (2023a) commitment-based approach or various neo-performative approaches in the spirit of Speas and Tenny (2003), require further assumptions. As such, self-talk serves as an ideal litmus test to probe into the syntactic structure at the very top. Finally, in as much as the analysis of the typology of self-talk is on the right track, we have new evidence for the hypothesis advocated in Wiltschko (2022), according to which language is for thought and communication. That is, the evidence from self-talk supports the view that interactional notions (such as speaker and addressee role as well as turn-taking management) are built into (universal) grammatical representations. Hence, we must conclude that interaction is built into our knowledge of language. But if self-talk is to be understood as an interactional phenomenon, which is regulated by interactional structure, then we also have to conclude that linguistic interaction (i.e., communication) cannot be defined as information exchange. Clearly, self-talk cannot be considered as such. However, as shown in Geurts (2018), if linguistic interaction is conceptualized as a way of negotiating commitments, then self-talk does not present an anomaly.

While the core goal of this paper was to introduce the phenomenon of self-talk as a fruitful empirical domain for linguistic analysis, I also submit that taking a linguistic perspective may shed light on the phenomenon itself. Recall from the discussion in Section 2.1 that self-talk is typically situated between thought and communication. It has in common with thought that it is private rather than social, and it has in common with communication that it is (or rather can be) overt, as in Table 1, repeated below for convenience.

Table 1: Self-talk lies between thought and communication.

thought	Self-talk	Communication
Private	Private	Social
Covert	Overt	Overt

Based on the linguistic evidence we have discussed, we must refine this view precisely because self-talk is not a unified phenomenon. What our results suggest is that the two criteria used in Table 1 are not adequate. First, while self-talk may be overt (unlike thought), it may also be covert. Second, while it is true that self-talk is not a social phenomenon in the sense that it does not involve the interaction between two or more people, this aspect does not allow for a distinction between the different modes of self-talk. Rather, what appears to be the defining criterion that allows for a distinction between the different modes of self-talk is the presence of an addressee role. But crucially, this is independent of whether there are two different people involved in the interaction (i.e., whether the interaction is social). This is summarized in Table 7.

Suppose, then, that the social dimension is linguistically irrelevant, at least for the phenomena we here consider. Similarly, the overt/covert distinction is also irrelevant in that it does not help to clearly distinguish between the different modes of talking: self-talk may be realized overtly or covertly. But if we are left with the presence of an addressee role as the defining criterion, we are led to conclude that *I*-centered self-talk is indistinguishable from thought while *you*-centered self-talk is indistinguishable from typical conversations, and thus communication. But if *I*-centered self-talk is indistinguishable from thought, then language itself is indistinguishable from thought. And this is, in fact, the view held by Vygotsky (1934/1986: 218), according to whom “[t]hought is not merely expressed in words; it comes into existence through them” (see also Kompa 2024 for a recent incarnation of this view). This is not to say that there are no forms of thinking that are independent of language, such as imagistic or affective processes, as well as representational and

Table 7: Thought and the different modes of talking.

	thought	<i>I</i> -centered self-talk	<i>you</i> -centered self-talk	Typical conversation
Social	✗	✗	✗	✓
Overt	✗	✗/✓	✗/✓	✓
Addressee	✗	✗	✓	✓

conceptual states, but propositional thought may, in fact, be only possible through language (see Hinzen and Sheehan 2015), if only in the form of inner speech. This view is also compatible with Deamer's (2021) view, according to which inner speech is necessary to bring our thoughts into consciousness. Thus, the linguistic characteristic of self-talk has the potential to shed light on the function of self-talk itself.

Before we conclude, a word of methodological caution is in order. Throughout this paper, I have made use of 1st and 2nd person pronouns to distinguish between *I*- and *you*-centered self-talk. This is however not necessary, and sentences used in self-talk do not require the use of such pronouns. To see this, consider the examples in (74), where the same sentence (*He is such an idiot*) is used in *I*-centered self-talk, as in (74a), and in *you*-centered self-talk (74b).

- (74) Alaka is at a dinner party where one of the guests is behaving idiotically.
 When Alaka goes to the bathroom and says to himself:
 a. I can't believe what he said. He is such an idiot.
 b. Alaka, do not start a conversation with him. He is such an idiot.

In isolation, it would be hard to discern whether this utterance instantiates a case of *I*-centered or *you*-centered self-talk. Though it is clear that such utterances may be part of either thinking out loud or having a conversation with oneself. Thus, neither the use of *I* nor *you* can be viewed as a necessary condition to identify the mode of self-talk. Moreover, the use of *I* is not a sufficient condition to identify the mode of self-talk characterized as thinking out loud. When having a conversation with oneself, the speaker can refer to them with *I* no matter whether the speaker is conceptualized as a disembodied voice addressing the person engaged in self-talk or whether it is the person engaged in self-talk addressing a reflection of themselves. It is for this reason that one can find self-talk examples that contain both 1st and 2nd person pronouns, as in (75).

- (75) Alaka is at a dinner party where one of the guests is behaving idiotically.
 When Alaka goes to the bathroom and says to himself:
 If I were you, I wouldn't start a conversation with him.

I assume that the presence of the 2nd person pronoun is a sufficient condition to identify this utterance as *you*-centered self talk (i.e., Alaka is having a conversation with oneself). I will have to leave the conditions of use for utterances with both *I* and *you* for future research.

In sum, I have shown that through a systematic exploration of the linguistic properties of self-talk, we may draw conclusions about the nature of grammatical representation of utterances, which in turn has implications for our understanding of the language faculty on the one hand and the nature of self-talk on the other. Regarding these conclusions, however, I suggest that a more in-depth investigation is

necessary. What I hope to have shown is that the linguistics of self-talk is a fruitful avenue of research, with implications for linguistic theory as well as the psychology of self-talk and its kin. The typology of the modes of talking laid out in this paper invites several new research questions, including the following.

(i) Cross-linguistic variation

We have seen that there is significant cross-linguistic variation regarding the use of pronouns in self-talk. Our analysis suggests a principled reason for this variation, namely the presence or absence of intrinsic social deixis in a given unit of language. Specifically, I have argued that social deixis in addressee-referencing elements is restricted to mirror-oriented *you*-centered self-talk. It remains to be seen whether this hypothesis holds up against further cross-linguistic examination. For example, the use of allocutive agreement has hardly been studied in the context of self-talk. The only study I am aware of is that of Alberdi (1996) for Basque. The same is true for languages with extensive honorific marking.

(ii) Child Development and Language Acquisition

As discussed in Section 2.1, the first serious studies of self-talk were based on children who go through a developmental stage where self-talk is frequent. This has led researchers to hypothesize that self-talk plays an important role in their cognitive and/or linguistic development. The linguistic profile of self-talk in children might shed new light on the question of whether self-talk serves in the development of externalizing thought (the Piagetian view) or whether self-talk serves in the development of internalizing linguistic interaction (the Vygotskian view). Specifically, we may explore whether children use *I*-centered or *you*-centered self-talk and if there is a correlation with other aspects of their language development and/or their cognitive development.

(iii) Neuro-diversity

Given the potential of shedding light on the relation between language, thought, and communication, it might be useful to explore the linguistics of self-talk in individuals with a neuro-diverse profile, such as autism-spectrum, schizophrenia, or aphasia. While there is a significant body of literature that studies self-talk (and inner speech) in neuro-diverse populations, to the best of my knowledge, the focus has not been on the linguistic profile of self-talk. Specifically, one might explore the use of *I*-centered versus *you*-centered self-talk in neuro-diverse populations and if there are correlations with other aspects of their cognitive profile.

(iv) The significance of non-canonical conversation

More generally, the linguistic profile of the modes of self-talk reveals the significance of exploring the linguistics of non-canonical conversations. If, indeed, classic sentence structure is embedded in layers(s) of structure that serve to regulate the

linguistic interaction and are thus sensitive to the identity of the interlocutors, this opens a new empirical domain of investigation. Specifically, we will need to control for the identity of the interlocutor in ways that allow for the construction and elicitation of minimal pairs. This is, of course, standard in studies that explore the linguistics of phenomena that are sensitive to the addressee (e.g., allocutive agreement, formal pronouns, honorific marking, and certain discourse particles). What the present study shows, however, is that conversations with non-canonical addressees are illuminating and provide striking evidence for the grammatical representation of the addressee role. I suggest that other types of non-canonical addressees (such as pets, infants, or machines) may serve to probe the range and limits of variation in the grammar of interaction. While these issues are typically assumed to fall in the domain of pragmatics, we have seen that in the case of self-talk, pragmatics alone cannot account for the observed linguistic facts. It remains to be seen whether similar conclusions can be drawn based on other types of non-canonical conversations.

I hope that with this target article, others will be inspired to explore some questions raised by the research reported here.

References

- Aijmer, Karin. 1987. *OH* and *AH* in English conversation. In Willem Meijs (ed.), *Corpus linguistics and beyond*. 61–86. Amsterdam: Rodopi.
- Alberdi, Xabier. 1996. *Euskararen tratamenduak: Erabilera* [Use of modes of address in Basque: Use]. Bilbao: University of the Basque Country PhD thesis.
- Alderson-Day, Ben & Charles Fernyhough. 2015. Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin* 141. 931–965.
- Alderson-Day, Ben, Susanne Weis, Simon McCarthy-Jones, Peter Moseley, David Smailes & Charles Fernyhough. 2016. The brain's conversation with itself: Neural substrates of dialogic inner speech. *Social Cognitive and Affective Neuroscience* 11. 110–120.
- Ambar, Manuela. 1999. Aspects of focus in Portuguese. In Laurice Tuller & Georges Rebuschi (eds.), *The grammar of focus*, 23–53. Amsterdam: John Benjamins.
- Anderson, Stephen. 1971. *On the linguistic status of the performative/constative distinction*. Bloomington, IN: Indiana University Press Linguistics Club.
- Ariel, Nana. 2022. Don't think before you speak: On the gradual formation of thoughts during speech. *Pedagogy, Culture and Society* 32. 361–373.
- Arppe, Antti & Juhani Järviö. 2007. Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3. 131–159.
- Austin, John. 1962. *How to do things with words*. Oxford: Clarendon.
- van der Auwera, Johan, Nina Dobrushina & Valentin Goussev. 2003. A semantic map for imperative-hortatives. In Dominique Willems, Bart Defrancq, Timothy Coleman & Dirk Noël (eds.), *Contrastive analysis in language*, 44–66. London: Palgrave.
- Banfield, Ann. 1982. *Unspeakable sentences*. London: Routledge & Kegan Paul.

- Barch, Deanna, Fred Sabb, Cameron Carter, Todd Braver, Douglas Noll & Jonathan Cohen. 1999. Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions. *NeuroImage* 10. 642–657.
- Bartels, Christine. 1999. *The intonation of English statements and questions: A compositional interpretation*. New York: Routledge.
- Bolinger, Dwight. 1989. *Intonation and its uses. Melody in grammar and discourse*. Stanford, CA: Stanford University Press.
- Brinthaup, Thomas, Scott Benson, Minsoo Kang & Zaver Moore. 2015. Assessing the accuracy of self-reported self-talk. *Frontiers in Psychology* 6. 570.
- Brinthaup, Thomas, Michael Hein & Tracey Kramer. 2009. The self-talk scale: Development, factor analysis, and validation. *Journal of Personality Assessment* 91. 82–92.
- Bunker, Linda, Jean Williams & Nate Zinsser. 1993. Cognitive techniques for improving performance and building confidence. In Jean M. Williams (ed.), *Applied sport psychology: Personal growth to peak performance*, 2nd edn., 225–242. Mountain View, CA: Mayfield.
- Burton, Strang & Lisa Matthewson. 2015. Targeted construction storyboards in semantic fieldwork. In Ryan Bochnak & Lisa Matthewson (eds.), *Methodologies in semantic fieldwork*, 135–156. Oxford: Oxford University Press.
- Carruthers, Peter. 2002. The cognitive functions of language. *The Behavioral and Brain Sciences* 25. 657–674.
- Chella, Antonio & Arianna Pipitone. 2020. A cognitive architecture for inner speech. *Cognitive Systems Research* 59. 287–292.
- Chomsky, Noam. 2017. Language architecture and its import for evolution. *Neuroscience and Biobehavioral Reviews* 81(B). 295–300.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads*. Oxford: Oxford University Press.
- Clark, Herbert. 1992. *Arenas of language use*. Chicago, IL: University of Chicago Press.
- Clark, Herbert. 1996. *Using language*. Cambridge: Cambridge University Press.
- Cohen, Antonie, René Collier & Johan 't. 1982. Declination: Construct or intrinsic feature of speech pitch? *Phonetica* 39. 254–273.
- Deamer, Felicity. 2021. Why do we talk to ourselves? *Review of Philosophy and Psychology* 12. 425–433.
- Diaz, Rafael. 1992. Methodological concerns with private speech. In Rafael Diaz & Laura Berk (eds.), *Private speech: From social interactions to self-regulation*, 55–81. New York: Lawrence Erlbaum.
- Duncan, Robert & Duncan Tarulli. 2009. On the persistence of private speech: Empirical and theoretical considerations. In Adam Winsler, Charles Fernyhough & Ignacio Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, 176–187. Cambridge: Cambridge University Press.
- Duncan, Robert & Allan Cheyne. 2001. Private speech in young adults. *Cognitive Development* 16. 889–906.
- Eckardt, Regine. 2020. Conjectural questions: The case of German verb-final Wohl questions. *Semantics and Pragmatics* 13(9). 1–54.
- Eckardt, Regine & Gisela Disselkamp. 2019. Self-addressed questions and indexicality – The case of Korean. *Sinn und Bedeutung* 23. 383–398.
- Egg, Markus and Malte Zimmermann. 2011. Stressed out! Accented discourse particles – the case of *doch*. *Sinn und Bedeutung* 16. 225–238.
- Etxepare, Ricardo. 1997. *The grammatical representation of speech events*. College Park, MD: University of Maryland PhD thesis.
- Farkas, Donka & Kim Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27. 81–118.

- Fernyhough, Charles. 2004. Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. *New Ideas in Psychology* 22. 49–68.
- Fernyhough, Charles. 2009. Dialogic thinking. In Adam Winsler, Charles Fernyhough & Ignacio Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, 42–52. Cambridge: Cambridge University Press.
- Fraser, Bruce. 1974. An examination of the performative analysis. *Paper in Linguistics* 7. 1–40.
- Friedman, Jane. 2013. Question-directed attitudes. *Philosophical Perspectives* 27. 145–174.
- Gaines, Robert. 1979. Doing by saying: Toward a theory of perlocution. *Quarterly Journal of Speech* 65. 207–127.
- Gärtner, Hans-Martin & Markus Steinbach. 2006. A skeptical note on the syntax of speech acts and point of view. In Patrick Brandt & Eric Fuß (eds.), *Form, structure, grammar*, 213–222. Berlin: Akademie-Verlag.
- Geurts, Bart. 2018. Making sense of self talk. *Review of Philosophy and Psychology* 9. 271–285.
- Giorgi, Alessandra. 2010. *About the speaker: Towards a syntax of indexicality*. Oxford: Oxford University Press.
- Goddard, Quinn, Elizabeth Ritter & Martina Wiltschko. 2022. Who am I talking to when I'm talking to myself? A cross-linguistic study. *Proceedings of the Annual Conference of the Canadian Linguistic Association*. <https://cla-acl.ca/actes/actes-2022-proceedings.html>.
- Goffman, Erving. 1981. *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.
- Goodhue, Daniel. 2024. Everything that rises must converge: Toward a unified account of inquisitive and assertive rising declaratives. In Daniel Goodhue, Manfred Krifka, True Trinh & Kazuko Yatsushiro (eds.), *Biased questions: Experimental results & theoretical modelling*. Berlin: Language Science Press.
- Hackfort, Dieter & Peter Schwenkmezger. 1993. Anxiety. In Robert Singer, Milledge Murphy & Keith Tennant (eds.), *Handbook of research on sport psychology*, 328–364. New York: Macmillan.
- Haddican, Bill. 2015. A note on Basque vocative clitics. In Beatriz Fernández & Pello Salaburu (eds.), *Ibon Sarasola Gorazarre*, 303–317. Bilbao: University of the Basque Country.
- Haddican, Bill. 2018. The syntax of Basque allocative clitics. *Glossa* 3(1). 101.
- Hale, Ken & Samuel Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.
- Hardy, James. 2006. Speaking clearly: A critical review of the self-talk literature. *Psychology of Sport and Exercise* 7. 81–97.
- Heim, Johannes. 2019. *Commitment and engagement: The role of intonation in deriving speech acts*. Vancouver: University of British Columbia PhD thesis.
- Heim, Johannes & Martina Wiltschko. 2020. Deconstructing questions: Reanalyzing a heterogeneous class of speech acts via commitment and engagement. *Scandinavian Studies in Language* 11. 56–82.
- Helmbrecht, Johannes. 2013. Politeness distinctions in pronouns. In Matthew Dryer & Martin Haspelmath (eds.), *The world Atlas of language Structures online*. Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/45> (accessed 21 June 2019).
- Hill, Virginia. 2007. Vocatives and the pragmatics-syntax interface. *Lingua* 117. 2077–2105.
- Hinzen, Wolfram & Michelle Sheehan. 2015. *The philosophy of universal grammar*. Oxford: Oxford University Press.
- Hockett, Charles. 1959. Animal “languages” and human language. *Human Biology* 31. 32–39.
- Holmberg, Anders. 2010. How to refer to yourself when talking to yourself. *Newcastle Working Papers in Linguistics* 16. 57–65.
- Hurlburt, Russell, Ben Alderson-Day, Simone Kühn & Charles Fernyhough. 2016. Exploring the ecological validity of thinking on demand: Neural correlates of elicited vs. spontaneously occurring inner speech. *PLoS One* 11(2). e0147932.

- Hurlburt, Russell, Christopher Heavey & Jason Kelsey. 2013. Toward a phenomenology of inner speaking. *Consciousness and Cognition* 22. 1477–1494.
- Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Jang, Youngyun. 1999. Two types of question and existential quantification. *Linguistics* 37. 847–869.
- Jang, Youngyun & Il-Kon Kim. 1998. Self-addressed questions and quantifier interpretation. *Korean Linguistics* 9. 191–209.
- Johns, Louise & Philip McGuire. 1999. Verbal self-monitoring and auditory hallucinations in schizophrenia. *The Lancet* 353. 469–470.
- Kitagawa, Chisato & Adrienne Lehrer. 1990. Impersonal uses of personal pronouns. *Journal of Pragmatics* 14. 739–759.
- Koguma, Tageshi, Katsunobu Izutsu & Yongtaek Kim. 2020. Monologic deixis: Two distinct conceptions behind reflexive speech event. *Proceedings of the 22nd Conference of the Pragmatics Society of Japan* 15. 169–176.
- Koguma, Tageshi & Katsunobu Izutsu. 2022. What's my name in absolute solitude? The essence of monologic selves in Japanese. *Studies of Language and Culture* 26. 19–31.
- Kompa, Nikola. 2024. Inner speech and pure thought—do we think in language? *Review of Philosophy and Psychology* 15. 645–662.
- Krifka, Manfred. 2013. Response particles as propositional anaphors. *SALT* 23. 1–18.
- Krifka, Manfred. 2015. Bias in Commitment Space Semantics: Declarative questions, negated questions, and question tags. *SALT* 25. 328–345.
- Krifka, Manfred. 2023a. Layers of assertive clauses: Propositions, judgements, commitments, acts. In Jutta Hartmann & Angelika Wöllstein (eds.), *Propositional arguments in cross-linguistic research: Theoretical and empirical issues*, 115–181. Tübingen: Narr.
- Krifka, Manfred. 2023b. Linguistik des Selbstgesprächs, mit Evidenz aus Schnitzlers “Leutnant Gustl”. *Grazer Linguistische Studien* 94. 343–365.
- Kuroda, Sige-Yuki. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis.
- Langsford, Steven, Amy Perfors, Andrew Hendrickson, Lauren Kennedy & Danielle Navarro. 2018. Quantifying sentence acceptability measures: Reliability, bias, and variability. *Glossa* 3(1). 37.
- Latinjak, Alexander, Alain Morin, Thomas Brinthaup, James Hardy, Antonis Hatzigeorgiadis, Philip Kendall, Christopher Neck, Emily Oliver, Małgorzata Puchalska-Wasył, Alla Tovaes & Adam Winsler. 2023. Self-talk: An interdisciplinary review and transdisciplinary model. *Review of General Psychology* 27. 355–386.
- Lee, Patricia. 1974. Perlocution and illocution. *Journal of English Linguistics* 8. 32–40.
- Leech, Geoffrey. 1976. Metalanguage, pragmatics and performatives. In Clea Rameh (ed.), *Semantics: Theory and application*, 81–88. Washington, DC: Georgetown University Press.
- Levinson, Stephen. 2019. Interactional foundations of language: The interaction engine hypothesis. In Hagoort, Peter (ed.), *Human language: From genes and brain to behavior*, 189–200. Cambridge, MA: MIT Press.
- Littell, Patrick, Lisa Matthewson & Tyler Peterson. 2010. On the semantics of conjectural questions. *UBC Working Papers in Linguistics* 28. 89–104.
- Lupyan, Gary & David Swingley. 2012. Self-directed speech affects visual search performance. *Quarterly Journal of Experimental Psychology* 65. 1068–1085.
- Malamud, Sophia. 2006. *Semantics and pragmatics of arbitrariness*. Philadelphia, PA: University of Pennsylvania PhD thesis.

- Marcu, Daniel. 2000. Perlocutions: The Achilles' heel of speech act theory. *Journal of Pragmatics* 32. 1719–1741.
- Mittwoch, Anita. 1976. Grammar and illocutionary force. *Lingua* 40. 21–42.
- Mittwoch, Anita. 1977. How to refer to one's own words: Speech-act modifying adverbials and the performative analysis. *Journal of Linguistics* 13. 177–189.
- Miyagawa, Shigeru. 2012. Agreements that occur mainly in the main clause. In: Lobke Aalbrecht, Liliane Haegeman & Rachel Nye (eds.), *Main clause phenomena. New Horizons*. 79–111. Amsterdam: John Benjamins.
- Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press.
- Miyagawa, Shigeru. 2022. *Syntax in the treetops*. Cambridge, MA: MIT Press.
- Miyagawa, Shigeru & Virginia Hill. 2023. Commitment phrase: Linking proposition to illocutionary force. *Linguistic Inquiry*. https://doi.org/10.1162/ling_a_00503.
- Murphy, Sean. 2015. I will proclaim myself what I am: Corpus stylistics and the language of Shakespeare's soliloquies. *Language and Literature* 24. 338–354.
- Nedergaard, Johanne & Gary Lupyan. 2024. Not everybody has an inner voice: Behavioral consequences of anendophasia. *Psychological Science* 35. 780–797.
- Oguro, Takeshi. 2017. Speech act phrase, conjectural questions, and hearer. *University of Pennsylvania Working Papers in Linguistics* 23. 191–199.
- Oyharzábal, Bernard. 1993. Verb agreement with non-arguments: On allocative agreement. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *Generative studies in Basque linguistics*, 89–114. Amsterdam: John Benjamins.
- Piaget, Jean. 1923/1962. *The language and thought of the child*. New York: Harcourt and Brace.
- Plato. 1970 (1892). Theatetus. The dialogues of Plato, 4. In Benjamin Jowett (ed.), *The Republic*, 107–280. Oxford: Oxford University Press.
- Portner, Paul, Miok Pak & Raffaella Zanuttini. 2019. The speaker-addressee relation at the syntax-semantics interface. *Language* 95. 1–36.
- Pratts, Jaydan, Gorana Pobric & Bo Yao. 2023. Bridging phenomenology and neural mechanisms of inner speech: ALE meta-analysis on egocentricity and spontaneity in a dual-mechanistic framework. *NeuroImage* 282. 120399.
- Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24. 657–720.
- Ritter, Elizabeth. 2024. Imperatives and prohibitives in Biblical Hebrew. In Ryan Bochnak, Eva Csipak, Lisa Matthewson, Marcin Morzycki & Daniel Reisinger (eds.), *The title of this volume is shorter than its contributions are allowed to be: Papers in honour of Hotze Rullmann*, 343–352. Vancouver, BC: UBCOPL.
- Ritter, Elizabeth & Martina Wiltschko. 2020. Interacting with vocatives. In *Proceedings of the Annual Conference of the Canadian Linguistic Association*. <https://cla-acl.ca/actes/actes-2020-proceedings.html>.
- Ritter, Elizabeth & Martina Wiltschko. 2021. Grammar constrains the way we talk to ourselves. In *Proceedings of the Annual Conference of the Canadian Linguistic Association*. <https://cla-acl.ca/actes/actes-2021-proceedings.html>.
- Ritter, Elizabeth & Martina Wiltschko. 2024. Pronouns beyond phi-features: The speaker-addressee relation in Japanese pronouns and its implications for formal pronouns. *Journal of Linguistics*. (online first). <https://doi.org/10.1017/S0022226724000306>.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: Handbook of generative syntax*, 281–337. Dordrecht: Kluwer.
- Ross, John. 1970. On declarative sentences. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 222–272. Waltham, MA: Ginn.

- Van Raalte, Judy, Allen Cornelius, Elizabeth Mullin, Britton Brewer, Erika Van Dyke, Alicia Johnson & Takehiro Iwatsuki. 2018. I will use declarative self-talk ... or Will I? Replication, extension, and meta-analyses. *The Sport Psychologist* 32. 16–25.
- Sacks, Harvey, Emanuel Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50. 696–735.
- Sadock, Jerry. 1969a. Hypersentences. *Research on Language and Social Interaction* 1. 283–370.
- Sadock, Jerry. 1969b. Super-hypersentences. *Papers in Linguistics* 1. 1–15.
- Sadock, Jerry & Arnold Zwicky. 1985. Speech act distinctions in syntax. In Timothy Shopen (ed.), *Language typology and syntactic description*, 155–196. Cambridge: Cambridge University Press.
- Sandsten, Karl, Dan Zahavi & Josef Parnas. 2022. Disorder of selfhood in schizophrenia: A symptom or a Gestalt? *Psychopathology* 55. 273–281.
- Schourup, Lawrence. 1982. *Common discourse particles in English conversation*. Columbus, OH: The Ohio State University PhD thesis.
- Schütze, Carson. 2016. *The empirical base of linguistics: Grammaticality judgments and Linguistic methodology*. Berlin: Language Sciences Press.
- Senay, Ibrahim, Dolores Albarraçin & Kenji Noguchi. 2010. Motivating goal-directed behavior through introspective self-talk: The role of the interrogative form of simple future tense. *Psychological Science* 21. 499–504.
- Sokolov, Aleksandr. 1975. *Inner speech and thought*. New York: Plenum.
- Son, Veronica, Ben Jackson, Robert Grove & Deborah Feltz. 2011. 'I am' versus 'we are': Effects of distinctive variants of self-talk on efficacy beliefs and motor performance. *Journal of Sports Sciences* 29. 1417–1424.
- Speas, Peggy & Carol Tenny. 2003. Configurational properties of point of view roles. In Anna Maria di Sciullo (ed.), *Asymmetry in grammar*, 315–344. Amsterdam: John Benjamins.
- Sprouse, Jon & Diogo Almeida. 2017. Setting the empirical record straight: Acceptability judgments appear to be reliable, robust, and replicable. *The Behavioral and Brain Sciences* 40. e311.
- Takubo, Yukinori. 2020. Nominal deixis in Japanese. In Wesley Jacobsen & Yukinobi Takubo (eds.), *Handbook of Japanese semantics and pragmatics*, 687–732. Berlin: Mouton de Gruyter.
- Theodorakis, Yannis, Robert Weinberg, Petros Natsis, Irini Douma & Panagiotis Kazakas. 2000. The effects of motivational versus instructional self-talk on improving motor performance. *Sport Psychologist* 14. 253–271.
- Thoma, Sonja. 2016. *Discourse particles and the syntax of discourse-evidence from Miesbach Bavarian*. Vancouver: University of British Columbia PhD thesis.
- Truckenbrodt, Hubert. 2004. Zur Strukturbedeutung von Interrogativsätzen. *Linguistische Berichte* 199. 313–350.
- Truckenbrodt, Hubert. 2006. On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32. 257–306.
- Vygotsky, Lev. 1934/1986. *Thought and language*. Cambridge, MA: MIT Press.
- Wechsler, Stephen & Larisa Zlatić. 2003. *The many faces of agreement*. Stanford, CA: CSLI Publications.
- Weigand, Edda. 2010. Language as dialogue. *Intercultural Pragmatics* 7. 505–515.
- Wiltschko, Martina. 2021. *The grammar of interactional language*. Cambridge: Cambridge University Press.
- Wiltschko, Martina. 2022. Language is for thought and communication. *Glossa* 7(1). <https://doi.org/10.16995/glossa.5786>.
- Wiltschko, Martina. 2024a. The meaning of *what*. In M. Ryan Bochnak, Eva Csipak, Lisa Matthewson, Marcin Morzycki & Daniel K. E. Reisinger (eds.), *The title of this volume is shorter than its contributions are allowed to be: Papers in honour of Hotze Rullmann*, 393–404. Vancouver, BC: UBCOPL.
- Wiltschko, Martina. 2024b. The syntax of talking heads. *Journal of Pragmatics* 232. 182–198.

- Wiltschko, Martina & Johannes Heim. 2016. The syntax of confirmational: A neo-performative analysis. In Gunter Kaltenböck, Evelien Keizer & Arne Lohmann (eds.), *Outside the clause: Form and function of extra-clausal constituents*, 305–340. Amsterdam: John Benjamins.
- Winsler, Adam. 2009. Still talking to ourselves after all these years: A review of current research on private speech. In Adam Winsler, Charles Fernyhough & Ignacio Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, 3–41. Cambridge: Cambridge University Press.
- Yetkin, Zerrin Thomas, Hammeke, Sara Swanson, George Morris, Wade Mueller, Timothy McAuliffe & Victor Haughton. 1995. A comparison of functional MR activation patterns during silent and audible language tasks. *American Journal of Neuroradiology* 16. 1087–1092.
- Zanutini, Raffaella. 2008. Encoding the addressee in the syntax: Evidence from English imperative subjects. *Natural Language and Linguistic Theory* 26. 185–218.
- Zimmermann, Malte. 2013. Ob-VL-interrogativsatz. In Jörg Meibauer, Markus Steinbach & Hans Altmann (eds.), *Satztypen des Deutschen*, 84–104. Berlin: de Gruyter.
- Zimmermann, Kathrin & Peter Brugger. 2013. Signed soliloquy: Visible private speech. *Journal of Deaf Studies and Deaf Education* 18. 261–270.
- Zu, Vera. 2015. Probing for conversation participants: The case of Jingpo. *CLS* 49. 379–389.
- Zu, Vera. 2017. *Discourse participants and the structural representation of the context*. New York: New York University PhD thesis.
- Zwicky, Arnold. 1974. Hey, what's your name. *CLS* 10. 787–801.