1 Interactional Language

It's not the language but the speaker that we want to understand.
Veda, Upanishads

1.1 Introducing Interactional Language

Language allows us to communicate things about the world we live in, how we perceive it, how we think of it, and how we evaluate it. Language allows us to gain insight into one another's mental worlds. There are thus two aspects of language: the role it plays in the thoughts about the world and the role it plays in the communication of those thoughts. This dual function sets up a dichotomy that has pervaded the study of language; it has defined different research agendas and methodologies. There are those traditions that take as their object of study the form of language used in the expression of thought, and there are those that take it to be its communicative function.

A formal (generative) linguist takes language to be a computational system that produces an infinite set of hierarchically structured expressions interpreted by our conceptual-intentional (C-I) system (Chomsky 2008). For the formal linguist, the object of study is humans' *competence* for language, rather than their *performance*. Exploring language competence makes necessary a particular methodology, unique to the generative enterprise: the elicitation of native speaker judgments.

A functional linguist takes language to be a means for communication; linguistic form is analyzed for its communicative functions. The distinction between competence and performance does not play a role and the normal way to collect data is by exploring the way people use language, that is, in natural settings.

A brief glance at language in interaction makes it clear that this dichotomy is spurious. Consider the interaction in (1) where $\mathscr{I}$ and $\mathscr{R}$ refer to the *initiating* and *reacting* roles, respectively.

(1) $\mathscr{I}$ Gal Gadot was amazing as Wonder Woman, eh?
$\mathscr{R}$ Yeah, I know, right?

$\mathscr{I}$ expresses their positive evaluation of Gadot's performance. $\mathscr{R}$ indicates that they agree. $\mathscr{I}$ and $\mathscr{R}$'s utterances contain more than the expression of these thoughts; there are several *units of language* (henceforth UoL) that contribute to managing the interaction, rather than adding content. The sentence-final particle *eh* signals that $\mathscr{I}$ assumes that $\mathscr{R}$ shares the same belief and encourages $\mathscr{R}$ to respond. Following Wiltschko and Heim (2016), I refer to such UoLs as *confirmationals*, as they are used to request confirmation.

In $\mathscr{R}$'s response, *yeah* indicates agreement; it doesn't seem to add much to the content of the utterance (*I know*). The use of sentence-final *right* appears odd: why ask $\mathscr{I}$ whether it is "right" that $\mathscr{R}$ knows? This is not something we typically need confirmation for. But in this context it makes the agreement more enthusiastic. The thoughts that are expressed in (1) (i.e., the propositional content) are simple, but the interaction conveys much more. Consider the same interaction without these UoLs.

(2) $\mathscr{I}$ Gal Gadot was amazing as Wonder Woman.
$\mathscr{R}$ I know.

The same thoughts are expressed, but the interaction has a distinctly different flavor. Unlike in (1), it is not clear whether $\mathscr{I}$ cares about $\mathscr{R}$'s opinion and $\mathscr{R}$'s response could be taken as rude; it seems to indicate that $\mathscr{I}$'s contribution is redundant. Thus, the sentence-peripheral UoLs change the quality but not the content of the exchange; they affect the *use* of language in interaction. Thus, a strict dichotomy between the form of language to express thought and the way it is used to convey these thoughts cannot be maintained: there are forms that affect the use of language. The forms that regulate interaction (i.e., use) have formal properties as well. Interestingly, these UoLs have received little attention in linguistics, in either formal or functional approaches.

The goal of this monograph is to fill this gap by exploring the formal properties of UoLs that regulate interaction, that is, the grammar of interactional language.

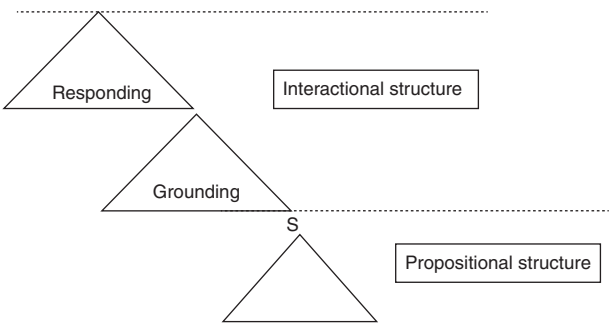
1.2 Toward a Grammar of Interactional Language

The core thesis I propose is that grammar not only configures the language used to convey thoughts, but also the language used to regulate interaction. Following generative assumptions, I take sentences to be hierarchically structured expressions, derived via a computational system. Utterances typically considered in grammatical analysis are sentences that convey thoughts about the world that can be true or false. I refer to the structure associated with such sentences as the *propositional structure*.

I propose that interactional language is derived by the same computational system and is therefore hierarchically structured: interactional structure dominates propositional structure. I argue that there are two core functions that characterize interactional structure: one serves to manage the common ground between the interlocutors: it is used to express things about mental worlds, rather than the world itself, and hence it plays an important role in the synchronization of minds. I refer to this function as the *grounding* function.

The second function concerns the management of the interaction itself (e.g., turn-taking). It aids the interplay between initiating and reacting moves. I refer to this function as the *responding* function. This is the core proposal I develop here: the *Interactional Spine Hypothesis* (henceforth ISH).

(3)



There are several arguments that interactional language is a part of grammar, in much the same way as propositional language is. I discuss and support each of these arguments in the course of this monograph. Interactional language is subject to well-formedness constraints; speakers have clear judgments about their use. This suggests that interactional language is part of competence. Interactional language shares much in common across languages, while displaying systematic variation in form, function, and distribution. It participates in paradigmatic contrasts and patterns of multi-functionality.

In addition to these properties of interactional language, I submit that assuming that there is a grammar of interactional language and that this grammar is essentially the same as the grammar of propositional language is really the null hypothesis. While it is true that the language of interaction is typically realized at the periphery of utterances, it still must be the case that its form and meaning are computed. Sentence-final particles must be combined with the host clause and with each other. I take it to be the null hypothesis that the same system that is responsible for computing propositional language is also responsible for computing interactional language. Moreover, the language of interaction is characterized by some of the same properties as propositional language. For example, it is often prosodically integrated with the propositional

structure and may combine with the same intonational tunes as propositional structure: sentence-final rise (indicated below as ↗) can be realized on a bare clause (4a) or on a confirmational (4b), but in the presence of a confirmational, it cannot be realized on the host clause no matter whether the confirmational also bears a final rise (4c) or not (4d).

(4) a. You have a new dog↗
b. You have a new dog, eh↗
c. *You have a new dog ↗, eh↗
d. *You have a new dog ↗, eh

This indicates that there must be a computational system responsible for regulating the realization of intonational tunes that cuts across the distinction between propositional and interactional language. The task at hand is to model how UoLs (including intonational tunes) combine with each other to derive the meaning and function of the complex utterance. This is precisely the role of syntax. I submit that a syntactic approach toward interactional language is not only possible, it can also be viewed as a necessary heuristic for exploring the grammar of interactional language. Since interactional language of this type is a novel empirical domain for formal typologies, we require a novel standard of comparison. This is precisely what the structure of interaction illustrated in (3) is meant to be. I take its success to be measured in terms of its empirical coverage. The goal of this monograph is to explore the linguistic reality of the structure of interactional language. If linguistic reality can be established, we can then go on to think about its psychological reality. It is a way of understanding the cognitive underpinnings of the interactional aspect of language, the mechanism that allows humans to synchronize their mental worlds through communicative interactions. In this way, the classic dichotomy between language as a means to express thought vs. language as a means to communicate such thoughts implodes.

1.3 The Significance of Interactional Language

There are several reasons to explore the grammar of interactional language. First, for the description of a language to be exhaustive, interactional language should be included. Strikingly, it is rarely mentioned in descriptive grammars of a language. Second, it is necessary for the sake of developing a typology: just like propositional language, interactional language is subject to variation. The question about the range and limits of variation is familiar from the point of view of propositional language, but it has not been systematically investigated for interactional language. The ultimate goal of finding out about language universals and variation is to find out about the cognitive underpinnings that underlie these universals; in the context of interactional language, this concerns

the cognitive underpinnings of our communicative competence, which are responsible for the logic of human verbal interaction.

A grammatical view on interactional language has its roots in a number of traditions (see Chapters 2 and 3). Recognizing the significance of interactional language was (in part) made possible by the recognition that language is typically embedded in an *act of speech.* Classic speech act theory (Austin 1962, Searle 1969) emphasizes that when we say things, we also do things: but what is the relation between what is said and what is done? The sentence itself is sometimes viewed as being enriched with meaning that regulates its use, namely force. For example, early on, Stenius (1967) argues that propositions by themselves are not units of communication; they need to be associated with illocutionary force. To capture its contribution, Stenius assumes that the sentences in (5) have the proposition (p) in common (the *sentence radical*), but that in addition, they combine with a *modal element* that signifies the force of the sentence (Stenius 1967 refers to this as mood). For an indicative, the proposition combines with an indicative modal element (I), for an imperative, there is an order (O), and for the interrogative, there is a question (?), as in (6).

(5) a. You live here now.
b. Live (you) here now!
c. Do you live here now?

(Stenius 1967: 254 (1–3))

(6) a. I(p) or p
b. O(p)
c. ?(p)

(Stenius 1967, 255 (1″–3″))

Similarly, Lewis (1970) suggests that non-declarative clauses may be analyzed as being embedded in a (covert) performative clause: the question in (7a) can be rendered as in (7b).

(7) a. Who is Sylvia?
b. I ask who Sylvia is.

If what we do with sentences (their force) is part of sentence meaning (as in (8), where p stands for proposition and F for Force), this aspect of interpretation is put squarely in the purview of grammar and is not just a matter of use in context.

(8) F(p)

On this view, certain aspects of sentence use are encoded within the sentence. This changes the way we think about the relation between form and meaning. In early structuralist theorizing, the object of study was patterns of form and how

they relate to meaning. It was a matter of understanding the relation between a particular form and its meaning (Saussure's program). This is also the focus in classic truth-conditional semantics (Frege's program). By adding conditions on its use (i.e., their force), the meaning of a sentence is enriched in ways that now involve the speaker and their intentions. The relation between the form of a sentence and its interpretation is mediated by the speaker's mind. Taking seriously the role of the speaker in the calculation of meaning goes hand-in-hand with recognizing the importance of what we do when we say things. The former is typically associated with Grice (1957) and the latter with Austin and Searle. Assuming something along the lines of (8) integrates use into meaning and thus transcends this distinction. But what aspects of use are encoded and how? What is the makeup of F in (8)?

Let us take as a starting point Lewis's (1970) propositional rendering of F, as in (7) (*I ask*): it includes reference to the speaker and a predicate of communication (*ask*). For most speech act theoretic approaches, these are the main ingredients of F: the speaker (and their intentions) and the particular type of speech act, its force.

However, speech acts do not occur in isolation; they are embedded in interaction. Thus, for the interpretation of speech acts, all interactants, the speaker *and* the addressee, are essential. According to Russell (1940: 204), there are three purposes of language: "(i) to indicate facts, (ii) to express the state of the speaker, (iii) to alter the state of the hearer."

Knowing who the addressee is, how they relate to the speaker, and what they know coming into the interaction affects what the speaker says and how they say it. And much of what we say comes with explicit instructions to the addressee as to what to do with what is being said. For example, some languages have dedicated mechanisms for indicating whether the addressee is higher or lower than the speaker on a social scale: this is reflected in forms of address, including the pronouns used to refer to the addressee. In addition, speakers are also sensitive to the knowledge states of their addressees. For example, in Bavarian German, the particle *fei* is used to indicate that the speaker believes that the addressee does not know p (Thoma 2016). *Fei* contrasts with *doch*, which signals that the speaker assumes that the addressee knows p (Thoma 2016).

(9) Martl is visiting Alex. Alex sets the dinner table for 2 and Martl assumes the second plate is for him. However, he has other plans that Alex doesn't know about.

a. I Hob **fei** koa Zeit zum Essn.
   I Have PRT no time to.DET eat
   'I don't have time to eat.'
b. *I hob **doch** koa Zeit zum Essn.

(10) Martl and Alex chitchat. Martl tells Alex he doesn't have time to stick around for dinner since he's going to the movies. Alex sets the dinner table for 2 and Martl assumes the second plate is for him.

a. *I hob **fei** koa Zeit zum Essn.
b. I hob **doch** koa Zeit zum Essn.

(Thoma 2016: 123 (9/10))

Finally, many of our utterances include a request for a response, for example in the form of rising intonation (see Heim 2019a for a recent discussion). Consider the difference between a declarative with falling and rising intonation, respectively.

(11) a. He has a new dog. ↘ He's so cute.
b. He has a new dog. ↗ *If so, what breed?

Falling intonation is compatible with the speaker keeping their turn; rising intonation is not. A rising declarative is explicitly requesting a response and hence the speaker must end their turn.

Thus, the addressee plays a role in the interpretation of speech acts; there are UoLs that are sensitive to the presence of an addressee. Thus, the addressee should be included in the makeup of F. The importance of the addressee is reflected in the classic speech act theoretic trichotomy (locution, illocution, perlocution).

In work that follows in the footsteps of classic speech act theory, however, the perlocutionary aspect is hardly addressed (see Chapter 3), and neither is the role of the addressee. The absence of the addressee is also evident in Lewis's (1970) propositional rendering of the speech act of questions in (7). The verb *ask*, which corresponds to the speech act, can also be used as a ditransitive verb (*I ask you* . . .). The propositional rendering of the illocutionary force, which includes both the subject and object of asking (*I ask you*), underlies the performative hypothesis of Ross (1970). He argues that every clause is dominated by a speech act structure that encodes this frame ($I\ V_{say}\ you$). Propositional content is embedded in structure that encodes the illocutionary force but is not spelled out. The insight behind Ross's analysis for declaratives is that even when we say things, we do things: we tell others about the world. These are the core ingredients of speech acts, according to Ross, and they are syntactically encoded (see Chapter 2). Ross's original proposal was abandoned shortly after publication, but the syntacticization of speech acts is currently an active research agenda. It has been made possible by the rise of functional projections that define the clausal architecture. And this has opened a new empirical domain to be described and analyzed: the language of interaction, such as the sentence-final particles introduced above. I show in this monograph that there

is a systematicity to the language of interaction, indicating that it makes use of similar building blocks as the propositional grammar it embeds. Specifically, it participates in patterns of contrast and patterns of multi-functionality, two of the hallmarks of universal grammatical categories (Wiltschko 2014).

There is a caveat, however. The broad type of information we are concerned with here, having to do with intentionality and interaction, can affect language in two ways: (i) it can be grammatically encoded, and (ii) it can come about via assumptions about the normal course of a conversation and the inferences that follow. It is the goal of this monograph to explore the grammatical underpinnings of interactional language.

The monograph is organized as follows. In Chapter 2, I introduce the body of research that aims to syntacticize speech acts. The core problem, I argue, is that it neglects the interactional nature of speech acts. This sets the stage for Chapter 3, where I review various frameworks that take seriously the interactional aspect of language. In Chapter 4, I introduce the ISH, its formal properties, and its methodological implications. Chapters 5 and 6 are the core empirical chapters: I explore the form and function of confirmationals and response markers. I show that the same formal mechanisms that serve to classify confirmationals also serve to classify response markers. This provides evidence for an underlying system that regulates the use of these markers: the interactional spine. In Chapter 7, I conclude and outline empirical and theoretical questions raised by the ISH with the intention to establish it as a research program.

2 The Syntacticization of Speech Acts

Ultimately, life is a chemical interaction. Heidi Hammel

2.1 Introduction

The goal of this monograph is to explore the grammar of interactional language. I argue that UoLs dedicated to regulating communicative interaction are part of syntactic structure. There are two core sources for this idea: (i) the syntacticization of speech acts, discussed in this chapter, and (ii) the development of speech act theory into a (dynamic) theory of interaction (Chapter 3). I attempt to combine these two lines of research and to explore the consequences. I show that existing models for the syntacticization of speech acts are missing an important aspect of language, namely its *interactional* component.

The chapter is organized as follows. I start with a brief introduction of classic speech act theory (section 2.2). I then discuss the relation between syntactic structure and speech acts. In section 2.3, I discuss approaches that take this relation to be a mapping relation: certain clause types are mapped onto certain speech acts via interpretive mapping rules. I then introduce and evaluate analyses according to which speech act structure itself is part of syntax (section 2.4). I introduce the initial instantiation of this idea, the so-called *performative hypothesis*, and I review arguments against it (section 2.5). But I also show that the syntacticization of speech acts can be upheld on the assumption that speech act structure is part of the functional architecture of natural language (section 2.6). I refer to this as the *neo-performative hypothesis*. I then argue that neo-performative hypotheses suffer from several weaknesses: most analyses fail to consider advances that have been made since classic speech act theory, namely the focus on the dynamic and interactional component of utterances. In section 2.7, I conclude.

2.2 Classic Speech Act Theory

When we talk to others, we not only say things, we also do things (Austin 1962). This insight triggered a large body of work exploring its consequences, both in philosophy and in linguistics.

2.2.1 Situating Speech Act Theory

At the time speech act theory was formulated, the dominant linguistic tradition was structuralism. It had introduced a focus on synchronic analysis and a distinction between *language as a system* (*langue*) and concrete instances of *language use* (*parole*). This echoes the division between language as a system for expressing thoughts and language as a means for communicating such thoughts. It also foreshadows the more cognitively oriented distinction between *competence* (what speakers know) and *performance* (what speakers do).

Taking a synchronic approach paves the way for exploring the relation between language and its context of use. But it still took a while for contextual information to become part of linguistic investigation. In structuralist traditions (including generative models), words and sentences are the units of analysis and within classic semantic traditions, Frege's principle of compositionality in (1) is the guiding principle for analysis.

(1) The meaning of a complex expression is determined by the meanings of its constituent expressions and the rules used to combine them.

It is mainly declarative denoting statements that are analyzed in this tradition; descriptions of facts about the world, which can be true or false. The focus on statements dates back to Aristotle, as does the significance of *truth*: "Not every sentence is a statement-making sentence, but only those in which there is truth or falsity" (Aristotle, *De interpretatione* (17a1–5), Edghill translation). The role of truth was formalized by Tarski in the 1930s. This in turn paved the way for truth-conditional definitions of meaning (Davidson 1967). Accordingly, to know the meaning of a sentence is to know its truth conditions, that is, what the world has to look like for the sentence to be true.

Against this backdrop, it comes as no surprise that the role of context did not receive much attention and neither did questions about what speakers intend to do when they utter a sentence. Within the Fregean tradition, the absence of this aspect of meaning follows from the fact that its goal was to develop a language that was adequate for logical argumentation rather than to understand the intricacies of natural language.

Nevertheless, even before Austin, we find approaches in which these notions played a role. For example, Ogden and Richards (1923) distinguish between *symbolic* and *emotive* meaning. Their notion of symbolic meaning corresponds