First saying, then believing: the pragmatic roots of folk psychology

Bart Geurts

ABSTRACT

Historical linguistics has revealed several pathways of language change that may guide our understanding of the evolution of mental-state attribution. In particular, it has been established that verbs of saying are often exapted for attributing a variety of mental states, including beliefs and intentions. For example, there are quite a few languages in which the literal translation of, "Boris said, 'I will win the elections'," may be used to convey that Boris thinks that he will win the elections or intends to win the elections. The objective of this paper is to analyse the pragmatic shifts underlying this pathway, and thus present the first articulate account of the evolution of belief/intention attribution.

1. INTRODUCTION

There are at least two social practices that, in our species, have reached an unprecedentedly high level of sophistication: communication and folk psychology. On the one hand, we are adept at getting things done and across by talking to each other, while on the other hand, we attribute all manner of mental states to one another and ourselves. It is commonly agreed that these practices are linked, and the currently received view is that the former builds on the latter. On this view, an utterance of "I'll be on time", for example, expresses that the speaker intends to be on time, and in order to achieve its purpose, the hearer must recognise that that is what the speaker means to convey. Hence, according to the received view, linguistic communication presupposes that speakers and hearers are able to attribute intentions to one another.

As is well known, this presupposition raises hairy issues concerning the development and evolution of human communication. These issues have been debated at length elsewhere,[1] and therefore I will keep this part of the story short.

1. See, e.g., Breheny (2006), Hirschfeld (2013), Thompson (2014), and Geurts (2019c) for a recent survey.

In a nutshell, then, if understanding a promise is primarily a matter of grasping the speaker's intentions, children must be able to attribute intentions before they can understand promises. But as it seems unlikely that two-year olds have this ability, it follows, contrary to fact, that two-year olds cannot understand promises. But then perhaps we should consider the possibility that two-year olds do have the ability to attribute intentions, after all. Unfortunately, however, prima facie this doesn't seem to be the case, there is no compelling evidence that it *is* the case, and there are no credible accounts of how this skill could be acquired within such a short period of time. Therefore, it will have to be supposed that this skill doesn't have to be learned, to begin with: it must be in our genes. Thus the ontogenetic problem is solved by passing it on to phylogeny, and the same questions that arose with respect to development now arise with respect to evolution.

Briefly, if it is a precondition for human communication that we are able to attribute intentions to one another, then it is a mystery where this ability comes from, and we cannot explain how human communication is possible in the first place. Therefore, I propose to consider the possibility that we can get much of our communicative business done without attributing mental states to each other. Elsewhere I have presented a pragmatic theory which allows for this possibility (Geurts 2018, 2019a,b), and in the following pages I will use that theory to explain how the social practice of attributing beliefs and intentions may have evolved once a distinctively human style of communication had been established. Starting from a social-communicative state which had advanced from that of non-human primates, but was still quite basic by sapient standards, my aim is to show how the practice of attributing beliefs and intentions may have evolved, without leaps or bounds, from certain discursive practices.

My focus is on beliefs and intentions for two main reasons. First, it is generally agreed that these are especially relevant for linguistic communication, and perhaps for human social interaction in general. Second, and more importantly, it is a lot harder to provide an evolutionary account for belief and intention attribution than for most other mental states. Compare pain, for example. Although most of us would be hard put to define what pain is, it has a distinctive and salient experiential quality, there are external circumstances that reliably tend to cause it, and pain has immediate external effects that are observable to others and oneself. This makes it relatively easy to see how we could spontaneously attribute pain to others. A week ago, I hit my thumb with a stone axe, thus causing a rather dreadful experience, wailing, and jumping-about. Today I see you hit your thumb with a stone axe, followed by wailing and jumping-about, and therefore I conclude that your experience must be the same as mine last week. How hard can it be?

The thing is that this will not work for all mental states, especially beliefs and intentions. Consider belief. On my way to work I collect lots of beliefs: that it is chilly; that there are children in the school playground; that the apple trees are in bloom; and so on. I register all these facts, and will be able to recall them later in the day, so they must count as beliefs. But I don't experience them as distinctive states of myself. None of these beliefs are likely to ever manifest themselves in any way or form, and while all of them are caused by visual perception, this is not something I notice. Much the same holds for intentions, and although they are generally at a lesser remove from overt action than beliefs, it is still the case that each of us has plenty of intentions that will not manifest themselves in the foreseeable future, if at all. Hence, although attributing beliefs or intentions may seem as easy as attributing pain, it is not nearly as easy to account for.

The main objective of this paper is to show how the folk-psychological practice of attributing beliefs and intentions may have emerged out of relatively basic discursive practices. My ambition is just that. I won't be arguing that these attributive practices couldn't or didn't evolve otherwise; nor that they evolved early, late, or neither. Most importantly, I won't have anything to say about beliefs and intentions, except as states we attribute to each other and ourselves; and that is how my discussion of these states is to be understood. My concern is with the folk psychology of beliefs and intentions, not their metaphysics, and although I incline to the view that beliefs and intentions exist for real, I won't presuppose that they do.

The general idea pursued in this paper is not new. It is that mental-state talk developed out of other, more basic uses of language. There is solid empirical evidence to support this idea. It is a well-attested observation, for example, that perception verbs like "see" and "hear" have often been extended to talk about epistemic states, as in English "I see what you mean" or "I hear what you're saying" (Sweetser 1990, Evans and Wilkins 2000). It is lesser known that, cross-linguistically, quotative expressions like "say" turn out to be closely related to mental-state expressions, and that there are quite a few languages in which quotatives are the main, if not the only, vehicle for discussing a wide variety of mental states, including beliefs and intentions (Pascual 2014). In such languages, the equivalent of the sentence "Betty said 'I'm pregnant'" may be used to convey that Betty thought she was pregnant or intended to get pregnant.

These are observations about documented and, in some cases, reconstructed languages of the last 5,000 years or so. But from here it is not a big leap to the idea that the same patterns may be used to explain the first origins of the attribution of beliefs, intentions, and other mental states. In a paper coeval with the current one, Moore (2019) proposes just that, but whereas Moore's main concern is to argue *that* mental-state attribution may have evolved out non-mentalist forms of language, mine is to present a detailed model of *how* the evolutionary

process may have unfolded, with special focus on the nexus between quotatives and mental-state attribution. The account I propose starts from the following hypothesis:

> *QA-hypothesis:* The practice of attributing beliefs and intentions evolved out of the quotative use of language.

This wording is not meant to imply that quotatives were the *only* source of belief/intention attribution. Moore points to the possibility that the practice of attributing epistemic states was an exapted use of perception verbs. That may be right, too, though this pathway will only account for the attribution of epistemic states. Be that as it may, it is not unlikely that folk psychology sprang from more than one source, but in the following I focus on a source that seems especially promising to me, since quotatives are known to have acquired an exceptionally wide range of folk-psychological uses.

The project of this paper is to flesh out the QA-hypothesis by providing a model of the pragmatic shifts underlying the pathway from quotation to the attribution of beliefs and intentions. The main objective of this model is to prove that it was possible for quotatives to gradually shift into uses that we would paraphrase in terms of belief and intention, but *without* assuming that the speakers who effected these shifts must have grasped what beliefs and intentions are before they started using quotatives to attribute them to one another.

I will proceed as follows. To begin with, I survey what I take to be the core features of our explanandum, which is the attribution of beliefs and intentions (§2). Next, I outline the pragmatic framework adopted in this paper, show how belief/intention attribution is accommodated by it (§3), and discuss the linguistic evidence supporting the QA-hypothesis (§4). Finally, I present my model of the evolutionary process underlying the transition from quotation to belief/intention attribution: starting from a baseline state in which our ancestors began to use some form of direct quotation, I describe a series of stages in which this form acquired new uses, until it came to be used for the attribution of beliefs and intentions (§5).

2. Aspects of mental-state attribution

Folk psychology is best seen as a aggregate of practices of which mental-state attribution is one (Astington and Baird 2005, Apperly 2011, Andrews 2012). But mental-state attribution is important, and as pointed out already, beliefs and intentions are especially important, socially as well as pragmatically. Here we are concerned with mental-state attribution as a discursive practice, which paradigmatically manifests itself in the use of verbs like “know”, “believe”, “intend”, and so on. It is sometimes suggested that this practice is of marginal significance,

at best, but the statistics suggest otherwise. In two major corpora of English, the British National Corpus (BNC, Burnard and Aston 1998) and the Corpus of Contemporary American English (COCA, Davies 2010), three mental-state verbs and two perception verbs rank among the 25 most frequently used verbs of English, with "know", "think" and "see" even being in the top 15:[2]

BNC: see (11), know (12), think (15), look (18), want (23)
COCA: know (10), think (12), see (14), want (17), look (18)

("Look" and "see" are included because perception, too, is one of the topics of our folk psychology.) Clearly, these are common verbs, whose frequencies are comparable with, e.g., "go", "take", or "get". These findings suggest rather strongly that explicit mental-state attribution is by no means a rare phenomenon. But what are the implications for the attribution of beliefs and intentions? The verb "believe" is not rare, but as belief is also expressed by way of "know" (knowledge entails belief) and many uses of "think" ("thinking of …" is an exception), we can infer that belief attribution is quite common. Similarly, although the verb "intend" *is* relatively rare, "want" often implies intention, as do the future tense ("I will …"), the intentional future ("I'm going to …"), and various other forms, which makes it likely that intention attribution, too, is common enough. Hence, there good reasons to doubt that explicit belief/intention is a marginal phenomenon.[3]

Our practices of attributing beliefs and intentions are rich and complex, and as my aim is to provide a partial model of the incipient stages of these practices, it is not to be expected that it will capture current practices in all their richness and complexity. Rather, my proposal aims at what I take to be the core features of belief/intention attribution. Briefly, these core features are: (1) beliefs and intentions are propositional attitudes; (2) they are associated with behavioural dispositions; (3) they are private; and (4) belief/intention attribution is a normative practice. In the remainder of this section I will take up these points in order.

1. *Propositional attitudes.* We take beliefs to have content which either agrees with the facts or not, and this content is referable seperately from the belief:

Fred thinks it's raining, which is true.

Here the "thing" that is true is the same as what Fred believes to be true. In philosophy, these "things" are commonly called "propositions". Propositions are a kind of content, or information, if you will, which is special in that it is true or

2. Word frequencies from ucrel.lancs.ac.uk/bncfreq (BNC) and wordfrequency.info (COCA).
3. Of course, these observations don't tell us anything about other languages, and it may be that the facts are different for other languages and cultures (Lillard 1998). Hence, empirical research might affect the generality of the argument I'm making here.

false. There are other kinds of content, like the content of the adjective "green", for example, which are neither true nor false; these kinds are non-propositional by definition.

In the first instance, to say that a belief is a propositional attitude is just to say that it involves truth-valued content. It is often supposed in the literature that propositions are structured objects and that having a belief requires that the owner has a mental representation of the proposition involved. Neither of these claims follows from the fact that beliefs are propositional attitudes. I doubt that our folk psychology has much, if anything, to say about propositional structure or mental representations, and therefore I will not assume that these are core features of our everyday practice of belief attribution. Our folk psychology treats beliefs as somethings that have truth-valued contents, but as far as I can tell it remains deeply non-committal on what kinds of things those contents are.

Intentions are propositional attitudes, too, and the bulk of the last two paragraphs applies to beliefs and intentions alike, the distinctive feature of the latter being that the propositional content of an intention specifies a goal of the owner. If Barney intends to go fishing, then it is Barney's goal to go fishing. But as goals are sometimes pursued unknowingly, and even unwillingly, having an intention entails but is not the same thing as having a goal.

2. *Dispositions*. We expect people's beliefs and intentions to affect their behavioural dispositions. If I attribute to Wilma the belief that it is cold outside, I expect her to behave accordingly, and likewise, if I attribute to her the intention of having lunch with me, then I expect her to act in such a way that she can keep our appointment.

To say that beliefs and intentions are supposed to affect their owners' behavioural dispositions is not to imply that a given belief or intention always has the same behavioural effects. Wilma's behavioural dispositions are shaped not only by her intention of having lunch with me, but also by other of her mental states. For example, if she thinks that we were to have lunch at Bonjour's, that's where she will try to be, whereas if she thinks we agreed on Ni Hao, she will be heading for Ni Hao. Therefore, in general, the relation between individual mental states and dispositions to act is not a function in the mathematical sense: different instances of a given mental state will tend to be associated with different dispositions to act.

However, there is an important exception. Speech acts are actions too, and our mental states also shape our discursive dispositions, the things we are disposed to say. Some of these things have a special relationship with our mental states. For example, if Fred believes that Wilma is in Bangkok, we expect him to be prepared to say that Wilma is in Bangkok, and vice versa, if Fred says that Wilma is in Bangkok, we expect him to believe that Wilma is in Bangkok. (These expec-

tations are only ceteris paribus, of course, but they are none the worse for that.) Likewise, if Fred intends to go to Singapore, we expect him to be prepared to say that he will go to Singapore, and vice versa, if Fred says he will go to Singapore, we expect him to intend to go to Singapore. In short, there is a quasi-equivalence between a "*x* believes …" and "*x* is disposed to say that …", as well as between "*x* intends to …" and "*x* is disposed to say that *x* will …"

3. *Privacy.* Although we usually act on the assumption that a person's actions reveal her mental states, we also concede that, at the end of the day, mental states are private; they are states one can feign being in. I can pretend to believe that Trump is exceptionally bright, for example, by saying that Trump is exceptionally bright. But pretending is never an absolute matter: even if we manage to hide our "true" mental states from others, they will reveal themselves in at least some of our private actions, for example, when we talk to ourselves or when we stop praying to a god that we no longer believe in.

If I say that Trump is exceptionally bright, there is a mismatch between what I say and what I believe, though we generally act on the assumption that such cases are the exception rather than the rule. Still, from an evolutionary perspective it is an important fact that people aren't always honest. It has often been suggested that this has been a driving force in the evolution of folk psychology: even if honesty is the norm, or taken to be the norm, the ability to make out what people "really" believe, want, etc., is likely to be fitness-enhancing.

The privacy of mental states is not to be confused with the notion that they are "internal" states. If taken literally, this notion entails that mental states must be located somewhere, and indeed mental states are often associated with various parts of the body. In many folk psychologies, mental states are between the ears, but occasionally they are placed in the heart, the liver, or the gut (whence "gut feelings"), and of course pains and itches are wherever they are felt, even if it be a missing limb. But wherever their residence, the presumed location of mental states seems to have little bearing on their functionality. In particular, it doesn't seem to be the case that different locations imply different predictions about the behavioural consequences of mental states. Basically, we don't seem to care very much where people keep their beliefs, as long as they have them; and beliefs are private wherever they may be (if they are anywhere at all).

The adjective "private" is polysemous: it has several related uses, which are not necessarily restricted to individuals, as witness the fact that we speak of "private conversations", "private meetings", and so on. The common core of these uses is that privacy presupposes a separation between an in-group who behave in certain ways (in the widest possible sense of "behave"), and an out-group who are not supposed to be party to the in-group's doings. In some cases, the in-group seeks to conceal their doings from the out-group, and they may go so

far as to act in such a way that the out-group is misinformed about their doings. Privacy in the individual, psychological sense may be seen as a special case, in which the in-group consists of a single person.

4. *Normativity.* Mental-state attribution is a normative practice: if the circumstances are so-and-so, you are *supposed* to be in such-and-such a mental state. You are supposed to empathise if your child is distressed; to be in pain if you break a finger; to know that you're driving a car if you're driving one; to intend to feed the cat if you promised to feed the cat; and so on. If such expectations aren't met, it's not just as if you turned out to be younger than I thought you were; rather, there is something *defective* about your state of mind. In some cases the defect is clearly moral. Unless you have a good excuse, you're blameworthy if you fail to empathise with your own child or don't intend to hold your promise. By contrast, there is nothing morally wrong about not knowing that you're driving, though you may be a liability on the road if you're in such a state.

Hence, mental-state attribution is associated with a variety of normative expectations, two of which are worth singling out. One is that mental-state attribution has a logic of sorts. If I attribute to you a certain set of mental states, then there will be other mental states that I feel entitled to attribute to you because I consider them to be implied by the first set. For example, if I take you to prefer Riesling to any other white, and to know that this glass contains Riesling while that glass does not, then I will expect you to prefer this glass to that one. Likewise, if I have evidence that you prefer Pinot gris to Riesling, and Riesling to Chardonnay, then I will expect you to prefer Pinot gris to Chardonnay. If you fail to meet such expectations, your preferences are inconsistent, and inconsistency is a vice, though not necessarily a major one.

Secondly, some of the mental states that are inferrable from what you say are normative, too. Certain conversational implicatures are of this type (Grice 1975, Geurts 2010, 2019a). For example, if I say that this is either Pinot gris or Riesling, then you are entitled to infer that I don't know which one it is. Sincerity inferences, which on some accounts are conversational implicatures (e.g. Geurts 2019a), also fall into this category. For example, if I tell you that this is Riesling, you're entitled to infer that I believe this to be Riesling.

To sum up: beliefs and intentions are propositional attitudes, they are private and dispositional, and the attribution of beliefs and intentions is a normative practice. For practical purposes as well as for the purposes of this paper, dispositionality is the key feature: attributing mental states to others and ourselves permits us to discern, predict, explain, and influence behavioural patterns. Therefore, it shouldn't come as a surprise that behavourial dispositions play an important role in the following; but the other features will be relevant, too.

3. COMMUNICATION AND COORDINATION

Being the hypersocial species that we are, we are in constant need of having to coordinate our actions so that we can play billiards, have meetings, go to the cinema together, build cars, live in the same block of flats, and so on. Our interactions require advance planning, and if others couldn't be counted upon to do their parts, action coordination would be impossible. That's why we make commitments: we commit ourselves to act as goalkeeper or umpire, to chair a session or give a talk, to wash or dry the dishes, and so on. Making commitments is a form of expectation management; a way of permitting others to rely on us to act in certain ways, so that they can coordinate their activities with ours.

While commitments sometimes come about implicitly, they may also be negotiated explicitly. That's what communication is for, and promises are the paradigm case. If Wilma promises Fred to walk the dog, she commits herself to walk the dog, and by the same token Fred becomes entitled to act on the assumption that Wilma will walk the dog. More generally, the chief purpose of discourse is to enable speakers to share commitments which help them to coordinate their actions.

Commitments have been used widely in the philosophy of language, rhetoric, speech act theory, and formal theories of dialogue.[4] On my account, commitment is a three-place relation between two individuals, x and y, and a propositional content, φ: x is committed to y to act on φ, or $C_{x,y}\varphi$ for short. So, if Wilma tells Fred, "I'll walk the dog", then as a result of Wilma's promise, $C_{w,f}[\![w$ will walk the dog$]\!]$, where $[\![w$ will walk the dog$]\!]$ is the proposition that Wilma will walk the dog.

To say that x is committed to y to act on φ is to say that x is committed to y to act in a way that is consistent with the truth of φ. I take this to entail that y is entitled by x to act on φ, and should y act on φ and φ turn out to be false, then y may hold x responsible for the consequences. Hence, commitment is a normative concept. It belongs to the same family of relations as obligation, duty, and responsibility, all of which are primarily directed towards others. That is, commitments are social relations first and foremost, not psychological states: x can be committed to act on φ without suspecting that he is thus committed, and indeed without entertaining the possibility that φ.

Wilma's commitment to act on $[\![w$ will walk the dog$]\!]$ is *telic:* she is committed make this proposition true. By contrast, if Wilma tells Fred, "I walked the dog" (past tense), she becomes committed to act on $[\![w$ walked the dog$]\!]$, and this commitment is *atelic:* the truth value of $[\![w$ walked the dog$]\!]$ was already

4. See, e.g., Hamblin (1971), Brandom (1994), Kibble (2006a,b), De Brabanter and Dendale (2008), Krifka (2015), Geurts (2019a,b).

decided before Wilma made her statement. Telic or atelic, Wilma's commitments constrain any acts she may consider, including her speech acts: once she has asserted that she walked the dog, she should not make any speech acts that entail or suggest that she didn't walk the dog, she should be prepared to reconfirm that she walked the dog, and so on. In fact, commitments often reveal themselves primarily in their owners' discursive dispositions.

Now suppose that, instead of promising Fred to walk the dog, Wilma says to *herself*: "I'll walk the dog." The theory outlined in the foregoing makes straightforward predictions about such self-directed speech acts: they will be associated with *self*-commitments of the form, $C_{x,x}\varphi$.[5] In this case, the self-commitment is telic: $C_{w,w}[\![w \text{ will walk the dog}]\!]$.

Like social commitments, self-commitments enable coordination. Wilma's promise to Fred entitles him to plan his activities on the assumption that Wilma will walk the dog, so he can take it as given that he won't have to do it, that Wilma will be out of the house for an hour, and so on. Likewise, Wilma's promise to herself is part of her own planning process, so she won't have to ask Fred to do it, she will be busy for at least an hour, and so on. In short, while Wilma's commitment to Fred enables Wilma and Fred to coordinate their actions, Wilma's commitment to herself enables her to coordinate her own actions. On this view, it is unsurprising that self talk is associated with a variety of higher cognitive functions, including reasoning, problem solving, planning and plan execution, attention, and motivation (Winsler 2009, Vicente and Martinez Manrique 2011).

Just as we make commitments to others so as to coordinate our actions with theirs, we make commitments to ourselves to coordinate our own actions. Thus commitment has two faces: a social and a non-social one. Given $C_{x,y}\varphi$, let's say that x's commitment is "social" if $x \neq y$, and that it is a "self-commitment" if $x = y$. It might be thought that social and self-commitments are very different beasts, but actually they serve the same purpose, i.e. action coordination; this is what unifies the two notions.

Social talk comes first, but once linguistic competence has started to develop, self talk doesn't lag far behind. Most normal-developing children begin to engage in self talk in their second or third year of life. Initially, self talk is fully overt and not always clearly distinct from social talk. Its use builds up until the fifth year, after which a slow process of internalisation sets in: self talk gradually becomes more truncated and harder for overhearers to follow, while more and more children report using inner speech. However, overt self talk never goes away entirely, and it remains in use throughout the lifespan.

5. This view on self talk is developed at greater length in Geurts (2018, 2019a). In those papers, I used the term "private commitment" instead of "self-commitment". I'm changing my terminology here, because I want to separate the privacy of mental states from the privacy of commitments. I do believe they are the same, but I don't want to presuppose that they are.

Communicating with ourselves comes naturally to us. We don't teach our children to talk to themselves, and even before they start speaking, children use a variety of gestures and points to share information with others, but will also gesture and point for themselves (Rodríguez and Palacios 2007, Delgado et al. 2009). Deaf children spontaneously sign for themselves just as hearing children talk to themselves (Kelman 2001, Gutierrez 2006). More remarkable still, self-address seems to come naturally to other species, as well: there is good evidence that non-human primates who have been taught the basics of a sign language are liable to spontaneously sign for themselves (Bodamer et al. 1994, Jensvold 2014). For example, Gardner and Gardner (1974) report that their foster chimpanzee Washoe was often seen "moving stealthily to a forbidden part of the yard signing 'quiet' to herself, or running pell-mell for the potty chair while signing 'hurry'." (p. 20)

The notion of self-commitment is crucial to my account, because it captures the common core of our folk notions of belief and intention: if Betty believes that she is pregnant, she is committed to herself to act on ⟦Betty is pregnant⟧; if she intends to become pregnant, she is committed to herself to act on ⟦Betty will be pregnant⟧. Self-commitments have all the basic features of folk mental states, as discussed in the previous section: self-commitments are propositional attitudes, they constrain our behavioural dispositions, they are normative, and they are private in the sense that they can be hidden from others. Hence, to attribute intentions and beliefs is to attribute self-commitments of the telic and atelic variety, respectively.

To sum up, the general picture of human communication that I propose is this. Social, other-directed talk is the basic form of discourse; self talk is a derived form. Social talk and self talk bring about social commitments and self-commitments, respectively, and the latter are either intentions (telic) or beliefs (atelic). On this view, the transition from social talk to self talk is predictably smooth and easy, and the primary function of self talk derives straightforwardly from the primary function of social talk.

In this section, I outlined a theory of how language is used to undertake commitments, which include beliefs and intentions, and what it means to have and act on commitments. The attribution of commitments (notably, beliefs and intentions) is another matter, to which we turn now.

4. Quotation and mental-state attribution

Quotation is meta-talk: talk about talk. It comes in two main varieties:

Wilma said: "I'm sick."	(direct quotation)
Wilma said (that) she was sick.	(indirect quotation)

Direct quotation is the most basic variety. It merely requires a reference to an individual x, a quotative expression, like the English verb "say", for example, and a copy of x's utterance. Someone who doesn't understand a word of Frisian may still be able to quote a question in Frisian:

She said: "Kinne jo my fertellen wêr't it húske is?"

By contrast, indirect quotation requires at least some grasp of both the grammar and the meaning of the reported utterance; if Wilma's "I'm sick" is quoted indirectly as "Wilma said she was sick", first person and present tense must be transposed into third person and past tense, respectively. Direct quotation is also more basic than indirect quotation in the sense that, whereas all languages provide grammatical means for quoting speech directly, not all languages have special devices for indirect quotation (though many of them do), and in oral languages that do distinguish between direct and indirect quotation, the former is generally preferred to the latter (Pascual 2014: 84-85).

Although verbs of saying are the most common quotation devices, they are not the only ones. Many language have so-called "evidentiality markers" to indicate hearsay; these markers tend to evolve out of verbs of saying (Aikhenvald 2004). Also, in many languages, including English, motion verbs are used to quote speech:

So she looks around and she goes: "Where is everybody?"

While quotation may appear to be a rather marginal use of language, it actually seems to be quite common. In English, for example, "say" ranks among the five most frequent verbs, which makes it more frequent than any mental-state verb (cf. §2), and in the Hebrew Bible, "say" is the most frequent verb of all (Wigram 1995, Sandler and Pascual 2019).[6] If these facts are anything to go by, meta-talk is rife.

Not only is quotation a basic and common use of language, it proves to be a remarkably versatile tool, too. In particular, as documented at length by Pascual (2014: chapter 4), quotatives are a common device for attributing all manner of mental states, including epistemic attitudes (belief, knowledge, etc.), pro-attitudes (urges, wants, intentions, etc.), and emotions. This observation is the empirical keystone of my account of the evolution of belief/intention attribution, and it holds across languages. In English, the connection between quotation

6. In the BNC corpus, "say" is the fifth most frequent verb; in the COCA corpus it ranks fourth.

and mental-state attribution is reflected in the grammar of mental-state reports, which mirrors that of speech reports:

Wilma said/thought: "I'm sick."
Wilma said/thought (that) she was sick.

Note, in particular, that the grammar of *direct* quotation is used to attribute thoughts. Reported self talk (or "verbal thought", as it is sometimes called) typically involves a reflexive construction:

Wilma said/thought to herself: "I'm sick."

Tellingly, in cases like this the choice of verb seems mainly stylistic, and the line between saying and thinking becomes quite thin.

Although the original purpose of quotative expressions is to refer to speech events, they often develop uses in which actual speech events are not implied at all. Pascual (2014: 94,96) gives the following examples from Spanish and Catalan:

Entonces al principio fue un poco de descoloque mental, DE DECIR: bueno, ¿qué está ocurriendo? (Spanish)
Lit. "So at first there was a bit of a mental confusion, OF SAYING: 'OK, what's going on here?'"

el dolor de no poder-te bellugar, dolor DE DIR no puc, no puc, no puc (Catalan)
Lit. "the pain of not being able to move, pain OF SAYING: 'I can't, I can't, I can't'"

In these examples, quotatives help to describe confusion and pain, respectively, and it is not implied that somebody actually said, "OK, what's going on here?", or, "I can't, I can't, I can't". Rather, the quotatives in these examples serve to indicate the kinds of state a person *might have* conveyed by uttering these expressions. If any speech is quoted at all, it is merely virtual, or "fictive", as Pascual calls it.

In the examples from Spanish and Catalan, quotatives merely play a supporting role in the attribution of mental states, since the states in question are tagged by psychological nouns (i.e. "descoloque mental" and "dolor", respectively). But there are also languages in which quotatives play the lead role in mental-state attribution. Many Australian Aboriginal languages, for example, feature a single quotative verb which variously translates as "say", "think", "want", and so on. A case in point is Warrwa, an extinct language formerly spoken in North-Western Australia:

kurrur-ngkay	ka-na-ngka-marra	ja-n	
black	will cook	SAYS	(McGregor 2007: 28)

Although this sentence can be interpreted as a direct quotation ("He says: 'I will make tea'"), McGregor reports that, when it was recorded, its intended meaning was unequivocally: "He wants to make tea"; no reference to a speech event was made or implied.

Another example is from Ungarinyin, which like Warrwa is an Aboriginal language spoken in North-Western Australia, though the two languages are not related:

ngurrba	nya$_2$-nga$_1$-yi-minda	a$_1$-ma	jirri	
hit	her$_2$ I$_1$ will take	he$_1$ SAY	he	(Spronck 2016: 259)

According to Spronck, this sentence can be used to convey a range of meanings, including: "He says: 'I will hit her'", "He thinks: 'I will hit her'", "He thinks that he will hit her", and "He wants to hit her".

Just to make sure that the phenomenon is not restricted to Australia, here's an example from Yurakaré, a nearly extinct language of central Bolivia (Pascual 2014: 92):

a-teshe	ti-yurujre	ku-ta-ja	otto-ja	mala-ti	ana	tumumu
sleeping	my owner	SAID	go out	go	that	frog

Pascual translates this as "Knowing that his owner was sleeping, the frog went outside and left", though I surmise that the same sentence might be used to attribute a belief.

Summing up, cross-linguistically quotative expressions are prone to acquire uses in which they support mental-state attribution. In some languages, at least some mental states can *only* be attributed by means of quotatives, because dedicated expressions for, e.g., beliefs or intentions are simply not available. In some languages, quotatives are the preferred means for referring to certain mental states, or they may be optional. English happens to belong to the last group.

Now it's time for a disclaimer. Thus far I have followed the linguistic literature in supposing that languages like Warrwa, Ungarinyin, and Yurakaré employ quotatives for attributing thoughts, beliefs, intentions, wishes, and other mental states. As a matter of fact, however, it is more accurate to say that these languages use quotatives for the attribution of what *we* would call "thoughts", "beliefs", "intentions", "wishes", and so on. This qualification may seem pedantic, but it is important to get this straight, for we are about to adopt an evolutionary perspective, and develop the QA-hypothesis, which says that the practice of attributing beliefs and intentions evolved out of the quotative use of language. This hypothesis is liable to be misread as presupposing that, before they started attributing beliefs and intentions, our ancestors already had an understanding of what these mental states are, on which their "mentalist" uses of quotatives were based. I don't want to presuppose any of that. My account is meant to be consistent with the possibility that our understanding of beliefs and intentions is shaped by our social practice of attributing these states. Granted, our ancestors may already have had some grasp of beliefs and intentions before they started engaging in discursive practices that we would describe in these terms. That is as it may be, but my account doesn't hinge on the premiss that it is true.

5. From quotation to belief/intention attribution

According to the QA-hypothesis, belief/intention attribution had its evolutionary roots in a special form of discourse, i.e. quotation, following a pathway that is still in use. Or, to put it the other way round, the QA-hypothesis takes a pattern of language change attested in historical time, and extrapolates it to prehistoric time. In the remainder of this paper I present an account of the evolutionary process underlying the transition from quotation to belief/intention attribution. This will involve: (*i*) identifying a baseline stage of discursive practices in which our ancestors began to use some form of direct quotation, and (*ii*) a series of stages in which this form acquired new uses, until (*iii*) it came to be used to attribute self-commitments, that is to say, beliefs and intentions.

Before we start developing this model, we will first consider how words are known to acquire new meanings, following principles of lexical change that have been documented by historical linguists. So we will first be concerned with the general principles underlying stage (*ii*).

Languages are always in flux, and one of the glories of linguistics is that, since the late 18th century, its understanding of how languages change has progressed enormously (see Campbell 2013 for a survey). As it turns out, language change is a regular process on all levels — phonological, morphological, syntactic, semantic — and many of the underlying regularities have been brought to light. For our purposes, the most relevant branch of historical linguistics is the study of lexical change, and here the central principle is quite straightforward: words are continuously exapted for uses that are closely related to uses they already have. For example, the word "orange" originally referred to a citrus fruit, and then came to be used to refer to its colour; many other nouns underwent the same object-to-colour shift, including "charcoal", "lime", "saffron", and so on. In these cases the relation between old and new uses is metonymical. In other cases it is analogical; a well-known example is the temporal uses of words that derive from locational/directional uses: "*around* noon", "*after* midnight", "*at* the beginning of March", "*towards* the end of April", and so on.

As these examples illustrate, new uses of a word need not replace existing ones. Usually old and new uses coexist at least for a while, and often enough for long periods of time; this is why the vast majority of words of any language are polysemous, which is to say that they have several related uses. Still, in many cases, derived uses eventually displace ancestral ones, as a result of which words may come to have uses that are far removed, if not opposite to, their original uses. For example, the adjective "silly" once meant "blessed" or "happy", and went through several intermediate uses before ending up meaning "foolish" or "stupid" (which, incidentally, were amongst the medieval uses of "nice").

According to the QA-hypothesis, at some point in prehistoric time our ancestors introduced a device for direct quotation, perhaps a verb like "say", which then spawned a cluster of related uses, some of which we would classify as belief/intention attribution. Our job is to characterise the baseline situation in which direct quotation was introduced, and then give a series of small and natural meaning shifts resulting in belief/intention attribution. We'll start with the baseline.

A very simple language will do. Direct quotation is just a matter of referring to a source, copying one of its utterances, and attributing the latter to the former. The most minimalistic way of achieving this is by pointing to a source, copying an utterance, and let the context do the rest. Hence, as far as linguistic requirements go, the developments we will be concerned with could have started quite early (which is not to say that they *did*). However, for expository convenience we will adopt a slightly more advanced baseline language featuring individual terms (like names and pronouns) and simple sentences of various types; basic English will be used as a proxy for this language.

Concerning the social setting, we will suppose that incipient normative practices had already been adopted. For example, we could imagine that property rights were supposed to be respected, promises supposed to be kept, that infringements of norms were sanctioned, and so on. Normative practices were necessarily embryonic at this stage, because more advanced forms of normativity typically require special linguistic devices, in particular meta-talk, to be implemented. If a promise is broken, for example, meta-talk is needed for taking the offender to task ("You promised to do the dishes!"), asking for justification or an apology, complaining to others, and so on. Since our baseline language lacks the means for "going meta", normative practices like these didn't exist, yet.

So much for the baseline. Now things start happening.

SAY$_0$

Into this baseline situation an expression for direct quotation is introduced, which we will represent by SAY. Small capitals will serve as a reminder that this is not to be thought of as modern English, and for the same reason SAY is not inflected. SAY$_n$ refers to a specific use of SAY, and SAY$_0$ is the primordial use, i.e. plain direct quotation. If x utters "y SAY$_0$ S", where x and y are individuals and S is a sentence, then x quotes y as uttering that very sentence. The quoted utterance will have been addressed to someone, z (which may but need not have been x), and have consequences for y's commitments, so these are conveyed as well:

By uttering "y SAY$_0$ S", x undertakes to act on the proposition that:
- y uttered S and
- y is therefore committed to some z to act on $[\![\text{S}]\!]$.

The commitment attributed to y has two main aspects: it has normative consequences for y and z, and it constrains y's behavioural dispositions. In general the second aspect is the most important for practical purposes, and this is doubly true in the primitive setting we are concerned with, because normative practices are as yet poorly developed. Therefore, in the following our focus will be mainly on the behavioural dispositions associated with y's commitment.

Since direct quotation is a linguistic universal (as we saw in the last section), it is a moral certainty that some form like SAY_0 will appear at some point. But what purpose could it serve, especially if it was introduced at an early stage of the evolution of language? My answer starts with the observation that humans take a keen interest in each other's doings. We are bent on coordinating our activities with one another, and since our actions are constrained by what we say, we pay attention even to utterances addressed to others. Direct quotation is a quite powerful tool, because it enables us to attend to utterances made out of earshot.

There are legion ways in which quoted talk contributes to shaping our behaviour. Quotation provides us with information that we wouldn't have had otherwise; it helps us assess other people's characters and attitudes; it enables groups to coordinate even if they aren't together; and so on. Presumably, it was mainly because of these benefits that quotation started to spread. There is another factor that may have played a role from the beginning, but anyway gained momentum with time: quotation was instrumental in going beyond the basic normativity that is our baseline. We have already seen that advanced normativity relies greatly on metalinguistic discourse. The point I'm making here is that even direct quotation, the most basic of metalinguistic devices, boosts the normative resources of its users. For example, in order to support his claim that there will be rain, x might say:

Our Great Chief SAY_0 "There will be rain."

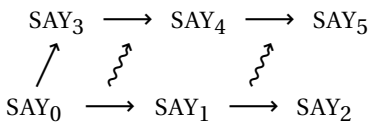
Or x might say, in order to convey that y's prediction was wrong:

You SAY_0 "There will be rain", but I haven't seen any.

Or she might use SAY_0 to report to her friends that y broke his promise and persuade them that sanctions are called for. Supporting a claim, refuting a prediction, calling for sanctions: these are all normative practices enabled by direct quotation, and the list is easily extended.

To sum up, direct quotation is a simple device, both linguistically and cognitively, while the same time it has great potential for benefiting and complicating social interactions. Therefore it may have emerged at a relatively early stage in the evolution of language and discourse.

In the following, I will describe five further uses of SAY, all of which derive from SAY_0 either directly or indirectly. These uses are related as follows:



Here the straight arrows represent direct derivations, and the squiggly arrows represent derivational analogies: the derivation of SAY_4 is analogous to, and supported by, the derivation of SAY_1, and likewise for SAY_5 and SAY_2. SAY_5 is belief/intention attribution, so according to the model, this use of SAY derives from SAY_0 via SAY_3 and SAY_4, and the derivation of SAY_4 and SAY_5 was supported by the derivation of SAY_1 and SAY_2. We'll continue our story with the last-mentioned uses of SAY.

SAY_1 *and* SAY_2

SAY_0 requires quotation to be strictly verbatim, so if Betty said, "Fred kissed Wilma", then her utterance would not license the statement, "Betty SAY_0 'Wilma was kissed by Fred'." It is only natural that speakers will be tempted to take liberties with this requirement, for the simple reason that, in general, content is of much greater interest than form. Thus, literalness is less urgent from the start, and SAY is bound to acquire uses which are less stringent with respect to the form of the quoted speaker's words:

> By uttering "y SAY_1 S", x undertakes to act on the proposition that:
> – y said something and
> – y is therefore committed to some z to act on $[\![\text{S}]\!]$.

This is still quotation, though in a looser sense than before: all that SAY_1 requires is that y said something which justifies the claim that she is committed to some z to act on $[\![\text{S}]\!]$.

Continuing to decrease in salience, the quotative element eventually drops out altogether:

> By uttering "y SAY_2 S", x undertakes to act on the proposition that y is committed to some z to act on $[\![\text{S}]\!]$.

Now it becomes possible to use SAY simply to attribute a commitment, without implying that it was associated with speech act. SAY_2 opens the floodgates for attributing a multitude of commitments that couldn't be attributed before. For example, if Barney is commonly known to have shot a bear, then he is supposed to be committed to act on $[\![$Barney shot a bear$]\!]$, and this may be expressed by "Barney SAY_2 there is a bear", regardless whether or not Barney said something that implied this commitment.

SAY$_3$

Thus far it was taken for granted that SAY$_0$, and likewise SAY$_1$ and SAY$_2$, were employed to attribute social commitments only (i.e. $y \neq z$). This assumption was mandated by the main objective of our project, which is to show how belief/intention attribution could evolve out of a social practice, namely, quoting other-directed speech. Therefore SAY$_0$, SAY$_1$, SAY$_2$ are stopping points along a cline from verbatim quotation of other-directed speech to plain attribution of other-directed commitments. But now imagine that SAY comes to be used for quoting *self*-directed speech:

> By uttering "y SAY$_3$ S", x undertakes to act on the proposition that:
> – y said S to y and
> – y is therefore committed to y to act on ⟦S⟧.

Like SAY$_1$, SAY$_3$ derives from SAY$_0$, but this time the derivation involves a transition from quoted social talk to quoted self talk. This is the key development in our narrative, and therefore we will dwell on it for a while.

As discussed in §3, the transition from social talk to self talk isn't a great leap.[7] Empirically, self talk comes naturally even to fledgling language users, like preschoolers and chimps with basic sign-language skills. Theoretically, the transition from social talk to self talk merely requires a shift from social commitments to self-commitments. But none of this implies that quoting self talk and attributing self-commitments are natural developments. Hence, the question we are faced with is this: once the practice of quoting social talk had been adopted, how could the practice of quoting self talk emerge out of it?

This is a bit of a concern, because unlike social talk, which we hear around us all day, overt self talk is sporadic. Although self talk it is a common enough phenomenon, it is progressively internalised from the first school years onwards. This is likely due, at least in part, to the social stigma attached to talking to oneself aloud (Duncan and Cheyne 2001, Duncan and Tarulli 2009). Whatever the reason, we relatively rarely hear others speak to themselves; so what prompted our ancestors to start quoting self talk?

My answer to that question consists of two parts, which differ greatly in substance. The less substantial of the two is that I see little reason to suppose that, in ancient times, overt self talk was stigmatised at all. Of course, it is impossible to establish whether this conjecture is correct, but the least we can say that, at the time when our ancestors began to engage in meta-talk, overt self talk may have been more common and more salient than it is now.

7. That is to say, it isn't a great leap within a commitment-based framework. In a mentalist framework it is hard to make sense of self talk, in the first place. See Geurts (2018) for discussion.

Be that as it may, the self talk we engage in is salient to *ourselves*, regardless if it is overt or not. Whereas beliefs and intentions are abstract, non-experiential states (§1), bouts of self talk are salient episodes in our experience, and their connections with subsequent behavioural dispositions are salient too. This is precisely why it may be useful to report self talk. Quoted self talk enhances speakers' linguistic resources in much the same way as the self-attribution of beliefs and intentions might have done, enabling them, for example, to come up with new forms of excuses ("I shot Fred because I said to myself, 'Lo, a bear!'") or to refer to mismatches, imagined or real, between social and private opinions ("Everybody cheered, but I said to myself, 'It's not over yet'").

Therefore, the transition from quoted social talk to quoted self talk need not have been such a great leap, and while the latter may be less obviously useful than the former, the potential benefits of quoted self talk are ample enough to warrant the assumption that the transition was going to be made eventually.

SAY$_4$ *and* SAY$_5$

Once verbatim quotation of self-talk was included in our ancestors' linguistic toolkit, it was bound to spawn non-verbatim uses just as we saw in the case of quoted social talk, and for the same reason: being more interested in content than in form, it was inevitable that speakers' focus would shift from the reported utterance to its effects. Thus, parallel to SAY$_1$ and SAY$_2$, and perhaps partly in imitation of these uses, the salience of the reported speech started waning (SAY$_4$) until it dropped out of the message altogether (SAY$_5$):

- By uttering "y SAY$_4$ S", x undertakes to act on the proposition that:
  - y said something to y and
  - y is therefore committed to y to act on $[\![\text{S}]\!]$.
- By uttering "y SAY$_5$ S", x undertakes to act on the proposition that y is committed to y to act on $[\![\text{S}]\!]$.

So there we are: what started as a linguistic expression for direct quotation is now used to attribute self-commitments, and depending on whether $[\![\text{S}]\!]$ represents a goal state of y or not, the attributed commitment will be an intention or a belief.

Starting its linguistic career as a device for verbatim quotation, prehistoric SAY has acquired five additional uses, which makes six in total. Modern English "say" retains at least four of these uses:

SAY$_0$: Donald said: "America has the best lawnmowers." [verbatim]
SAY$_1$: Vladimir said: "Russia has the best lawnmowers." [in Russian]
SAY$_3$: Donald said to himself: "America has the best lawnmowers." [verbatim]
SAY$_4$: Vladimir said to himself: "Russia has the best lawnmowers." [in Russian]

The remaining uses are harder to find for “say”, which is probably due to fact that English has dedicated expressions for these, like “be committed”, “think”, “know”, and so on.

6. Conclusion

The communicative resources that the great apes have at their disposal are paltry compared to human languages, which are stunningly complex on various levels: phonetic, phonological, morphological, syntactic, semantic; and the myriad connections between these levels add to the complexity of the system as a whole. The driving force behind this complexity is pragmatic: whereas chimpanzee gestures have a quite restricted range of uses, language is used in so many ways that counting them is pointless. The key innovation in our lineage was that, once upon a prehistoric time, hominins began to communicate so as coordinate their actions beyond the immediate future. That’s when communication started to become commitment-based.

When did folk psychology enter the scene? In order to answer that question, we need to refine it first. If “folk psychology” refers to a set of capacities, then some of these, e.g. the ability to recognise arousal or pain, may have deep evolutionary roots indeed. But if folk psychology is seen as a social practice, as it is here, then it is part and parcel of our communicative practices, and it is only to be expected that it requires non-trivial communicative resources.

Historical linguistics has established several pathways of language change that may guide our understanding of mental-state attribution, socially understood. In particular, it is an undisputed fact that verbs of saying are often exapted for attributing a variety of mental states, including beliefs and intentions (Pascual 2014), and that verbs of seeing and hearing are often exapted for epistemic uses (Evans and Wilkins 2000). My project in this paper was to develop the first pathway from an evolutionary perspective; to the best of my knowledge, this is the first articulate account of the evolution of belief/intention attribution. The second pathway may be developed along somewhat similar lines, though the details are different and remain to be worked out.

REFERENCES

Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford University Press.

Andrews, K. (2012). *Do apes read minds? Toward a new folk psychology*. Cambridge, Massachusetts: MIT Press.

Apperly, I. A. (2011). *Mindreaders: the cognitive basis of "theory of mind"*. Hove: Psychology Press.

Astington, J. W. and J. A. Baird (2005). Introduction: Why language matters. In: J. W. Astington and J. A. Baird (eds.), *Why language matters for theory of mind*, pp. 3–25. Oxford University Press.

Bodamer, M., D. Fouts, R. Fouts, and M. Jensvold (1994). Functional analysis of chimpanzee (pan troglodytes) private signing. *Human evolution* 9: 281–296.

Brandom, R. (1994). *Making it explicit*. Cambridge, Massachusetts: Harvard University Press.

Breheny, R. (2006). Communication and folk psychology. *Mind and language* 21: 74–107.

Burnard, L. and G. Aston (1998). *The BNC handbook: exploring the British National Corpus*. Edinburgh University Press.

Campbell, L. (2013). *Historical linguistics: an introduction* (third edition). Edinburgh University Press.

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of english. *Literary and linguistic computing* 25: 447–464.

De Brabanter, P. and P. Dendale (2008). Commitment: the term and the notions. In: P. De Brabanter and P. Dendale (eds.), *Commitment*, pp. 1–14. Amsterdam: Benjamins.

Delgado, B., J. C. Gómez, and E. Sarriá (2009). Private pointing and private speech: developing parallelisms. In: A. Winsler, C. Fernyhough, and I. Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, pp. 153–162. Cambridge University Press.

Duncan, R. and D. Tarulli (2009). On the persistence of private speech: empirical and theoretical considerations. In: A. Winsler, C. Fernyhough, and I. Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, pp. 176–187. Cambridge University Press.

Duncan, R. M. and J. A. Cheyne (2001). Private speech in young adults: task difficulty, self-regulation, and psychological predication. *Cognitive development* 16: 889–906.

Evans, N. and D. Wilkins (2000). The semantic extensions of perception verbs in Australian languages. *Language* 76: 546–592.

Gardner, B. and R. Gardner (1974). Comparing the early utterances of child and chimpanzee. In: *Minnesota symposium on child psychology*, Volume 8, pp. 3–23. Minneapolis: Minnesota University Press.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Geurts, B. (2018). Making sense of self talk. *Review of philosophy and psychology* 9: 271–285.

Geurts, B. (2019a). Communication as commitment sharing: speech acts, implicatures, common ground. *Theoretical linguistics* 45: 1–30.

Geurts, B. (2019b). Commitment continued. *Theoretical linguistics* 45: 111–125.
Geurts, B. (2019c). What's wrong with Gricean pragmatics. In: *ExLing 2019: Proceedings of the 10th international conference of Experimental Linguistics*, Lisbon.
Grice, P. (1975). Logic and conversation. In: P. Cole and J. Morgan (eds.), *Syntax and semantics 3: Speech acts*, pp. 41–58. New York: Academic Press.
Gutierrez, E. (2006). The self-regulatory function of sign language. In: I. Montero (ed.), *Current research trends in private speech*. Madrid: Publicaciones de la Universidad Autónoma de Madrid.
Hamblin, C. L. (1971). Mathematical models of dialogue. *Theoria* 37: 130–155.
Hirschfeld, L. A. (2013). The myth of mentalizing and the primacy of folk sociology. In: M. Banaji and S. A. Gelman (eds.), *Navigating the social world: what Infants and other species can teach us*, pp. 101–106. Oxford University Press.
Jensvold, M. L. (2014). Experimental conversations: sign language studies with chimpanzees. In: M. Pina and N. Gontier (eds.), *The evolution of social communication in primates: a multidisciplinary approach*, pp. 63–82. Springer.
Kelman, C. (2001). Egocentric language in deaf children. *American annals of the deaf* 146: 271–279.
Kibble, R. (2006a). Reasoning about propositional commitments in dialogue. *Research on language and computation* 4: 179–202.
Kibble, R. (2006b). Speech acts, commitment and multi-agent communication. *Computational and mathematical organization theory* 12: 127–145.
Krifka, M. (2015). Bias in commitment space semantics: declarative questions, negated questions, and question tags. In: S. D'Antonio, M. Moroney, and C. R. Little (eds.), *Proceedings of SALT 25*, Washington, pp. 328–345. LSA Open Journal Systems.
Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological bulletin* 123: 3–32.
McGregor, W. (2007). A desirative complement construction in Warrwa. In: J. Siegel, J. Lynch, and D. Eades (eds.), *Language description, history and development: linguistic Indulgence in memory of Terry Crowley*, pp. 27–40. Amsterdam: John Benjamins.
Moore, R. (2019). The cultural invention of mind-modelling. Submitted for publication.
Pascual, E. (2014). *Fictive interaction: the conversation frame in thought, language, and discourse*. Amsterdam/Philadelphia: Benjamins.
Rodríguez, C. and P. Palacios (2007). Do private gestures have a self-regulatory function? A case study. *Infant behavior and development* 30: 180–194.
Sandler, S. and E. Pascual (2019). In the beginning there was conversation: fictive direct speech in the Hebrew Bible. In: C. Gentens, M. S. Sansiñena, S. Spronck, and A. Van Linden (eds.), *Irregular perspective shifts and perspective persistence: discourse-oriented and theoretical approaches*, pp. 250–276. Benjamins.
Spronck, S. (2016). Evidential fictive interaction (in Ungarinyin and Russian). In: E. Pascual and S. Sandler (eds.), *The conversation frame: forms and functions of fictive interaction*, pp. 255–275. Amsterdam/Philadelphia: John Benjamins.
Sweetser, E. E. (1990). *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*. Cambridge University Press.

Thompson, J. R. (2014). Meaning and mindreading. *Mind and language* 29: 167–200.
Vicente, A. and F. Martinez Manrique (2011). Inner speech: nature and functions. *Philosophy compass* 6: 209–219.
Wigram, G. V. (1995). *The Englishman's Hebrew concordance of the Old Testament: coded with Strong's concordance numbers*. Peabody, Massachusetts: Hendrickson.
Winsler, A. (2009). Still talking to ourselves after all these years: a review of current research on private speech. In: A. Winsler, C. Fernyhough, and I. Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*, pp. 3–41. Cambridge University Press.