Reference Across Pathologies: A New Linguistic Lens on Disorders of Thought

Abstract: According to a linguistic tradition here termed ‘Cartesian’, language is relegated to an expressive system considered to provide the means to encode or communicate an independently constituted thought process. An alternative vision here termed ‘un-Cartesian’ regards language as an organizational principle of human-specific thought, with the implication that thought of the same type would not become available to a cognitive system without language and that clinical thought disturbances implicate language dysfunction. I here explore the latter view in the context of intra-species variation of the human cognitive type: cognitive disorders that, as in the case of autism and schizophrenia, come with language-related clinical symptoms. If language is the configurator of human-specific thought, cognitive and linguistic phenotypes should illuminate one another. I specifically review evidence for impairment in one universal linguistic function, namely reference. Linguistic meaning is referential meaning: we cannot utter sentences without referring to persons, objects, and events, based on lexica-lized concepts that provide descriptions of these referents. Reference in this sense takes a number of human-specific forms, from generic to specific, deictic and personal ones, which empirically co-vary with forms of grammatical organization. As reference in some of these forms proves to be highly vulnerable across major mental disorders, grammar is thereby linked to forms of thought and selfhood critical to normal cognitive functioning. In this way clinical linguistic and cognitive diversity provides an important new window into the foundational question of the thought-language relationship and the cognitive significance of grammar.

1 Introduction

Grammatical organization is universal in human populations, which naturally raises the question of its biological basis and universal principles. Pursuing this...
question, the universal grammar project in its twentieth century shape (Chomsky 1965) has largely abstracted from the connections between language and thought, focusing on language-specific principles hypothesized to define an invariant, genetically determined language faculty. These principles are taken to restrict the scope of possible linguistic diversity in human populations that are genetically and cognitively uniform. Insofar as clinical genetic and cognitive diversity has been brought to bear on universal grammar research, the focus has been on conditions isolating language from non-linguistic cognition as well. In particular, aphasia following lesions in left perisylvian cortex is language-specific by definition. Clinical impression here moreover often tends to suggest, not that the thought system becomes pathological, but that patients struggle to get essentially normal thoughts across. Dissociations between language and reasoning in various cognitive domains (Varley 2014), as well as differences in their respective neural correlates (Fedorenko and Varley 2016), have been taken to further support a ‘modular’ viewpoint separating the neural basis of ‘language’ from that of ‘thought’. Specific Language Impairment (SLI), too, is specific to language by name and definition.¹

A number of other clinical conditions, however, such as autism spectrum disorder (ASD), Alzheimer’s disease (AD), Huntington’s disease (HD), or schizophrenia (SZ), systematically implicate language changes along with disorders of thought.² These have comparatively received much less linguistic attention, just as, in turn, little neurological and especially neuropsychiatric attention has turned to language as a neurocognitive variable potentially mediating human-specific forms of rationality and mental health. This neglect is not an accident but at least in part a reflection of our prior foundational conceptions of what language is and what role it plays in human cognition. One such conception is that language is an expressive system dedicated to the communication of the human thought process to others, while not being inherent to this process itself. If in this case the thought process disintegrated clinically, we would naturally seek to identify the non-linguistic neuro-cognitive variables (e.g. memory,

¹ Neither aphasia nor SLI will be treated here in any detail, though both are significant to the un-Cartesian research program outlined below. Note, though, that how aphasia affects cognition is an open question (see e.g. Baldo et al. 2010; for evidence of cognitive impairment; see also Zimmerer et al. 2016). Classical putative evidence for modularity from both Williams syndrome and SLI (Pinker 1994) has also by now been questioned (Mervis and Becerra 2007; Laing et al. 2002; Karmiloff-Smith et al. 2003; Davies et al. 2016; Norbury et al. 2014).

² Neither AD nor HD will be treated here in any detail either. For recent evidence that language is a fine-grained measure of cognitive change and disease progression in these cases as well, see Ahmed et al. (2013) and Pekkala et al. (2013) for AD, and Hinzen et al. (2017) for HD.
attention, executive functioning, etc.) that might account for this change, an approach that has defined the field of cognitive neuropsychiatry (Frith 1992).

I will here call this expressive view broadly and technically, and partially for dialectical purposes, the ‘Cartesian’ one, and contrast it with an ‘un-Cartesian’ one. The term ‘Cartesian’ is chosen to allude to eighteenth century ‘Cartesian linguistics’ as characterized in Chomsky (1966), which posits an independent universal thought system as structured by logic, which grammar, insofar as it is rational, ‘mirrors’; but also to twentieth century generative linguistics, which separates the study of language and of thought (Chomsky 1965) (for historical remarks see Hinzen 2012; Hinzen and Sheehan 2013: Chapter 1). Functional and cognitive linguists, too, however, regard language as rooted in principles of human cognition that are not in turn conditioned by language, thereby breaking the connection between language and human-specific thought (see e.g. Tomasello 2008). Cartesian in my present terms is also the view that language is not merely a communicative device translating a private code into a public medium, but a ‘supra-communicative’ one that can enhance thought, be it through the provision of words viewed as cognitive ‘artefacts’ or ‘vehicles’ (Clark 1998), or by expressing it as ‘inner speech’ and making us conscious of it (Carruthers 1996, 2002). Of course the Cartesian view also centrally includes the traditional hypothesis that thought is coded in a distinct ‘language of thought’ (LoT) (Fodor 1975; Carruthers 2002), often considered to structure the minds of linguistic and non-linguistic species alike.

Meanwhile in recent generative linguistics, the question of the ‘interface’ between two systems identified as ‘language’ and ‘thought’, respectively, has moved center-stage (Chomsky 1995). The hope has been that principles of the organization of the ‘computational system’ of language (on the ‘language’ side of the posited interface) will in part derive from the way it is embedded in non-linguistic thought systems – the so-called ‘conceptual-intentional’ (C-I) systems of Chomsky (1995) posited on the other side of this interface (illustrating another version of a Cartesian divide). These systems are assumed to be given prior to language evolution in hominins and to have constrained the design of the language faculty externally, with different views on how ‘tight’ this interface is (Chomsky 2007; Jackendoff 2002). In Chomsky (2007), Chomsky identifies a different view said to be ‘more radical’, according to which there is no interface (Hinzen 2006, 2007). Grammar is the organizational principle of the sapiens-specific thought process, which on this view would not exist without speech (or sign). The language-thought dichotomy (for this particular type of thought) thus collapses, thereby erasing a standard chicken-and-egg problem as well. This latter view will here be called ‘un-Cartesian’ and be developed in Section 2 as it has evolved over the last decade (via Hinzen 2007; Sheehan and Hinzen 2011;
Both Cartesian and un-Cartesian views of language can be empirically explored in the context of cognitive disorders, which is the subject of Section 3. If language defines a species-specific cognitive type, then cognitive and linguistic diversity in our species should match in the sense that their respective forms of diversity can be meaningfully and systematically mapped onto each other. Depending on what cognitive dysfunction we see and which cognitive functions language is hypothesized to mediate, linguistic and cognitive types should align. It is clear that neither ASD nor SZ can have complete animal models since language is intrinsic to core aspects of their clinical presentations. Yet the exact ways in which they implicate language has remained unclear and contested. Here the un-Cartesian approach makes a specific prediction based on its central claim that a species-specific form of reference is a core cognitive function of language in humans. Comparative evidence has provided evidence that this particular form of reference is discontinuous with forms of reference in non-grammatical beings, such as alarm call systems in monkeys (Cheney and Seyfarth 1990). These have been suggested to resemble innate human vocalizations like crying, grunting, sobbing or laughter more than speech (Deacon 2006). In particular, both are confined to relatively specific constellations of vocalizations, emotional states, and stereotypic referential contexts. The un-Cartesian research hypothesis is that grammatical organization frees reference from these constellations and bases it on lexicalized concepts that can be freely retrieved and combined independently of external stimulation. This allows cognitive flexibility and creative thought, which can be socially shared due to its linguistic format, representing the beginnings of the idea of an objective world. The core claim of Section 3 will be that available evidence from both ASD and SZ support a link between reference in this sense and core symptoms of these disorders.

Linguistic and cognitive diversity in clinical populations thus informs linguistic theory, by illuminating the foundational question of what language is. In turn, linguistics can shed new light on cognitive disorders, potentially evolving towards a ‘science of cognition’ (i.e. the human cognitive type) rather than

---

3 Darwin could be seen as falling on the un-Cartesian side of this divide: ‘If it be maintained that certain powers, such as abstraction, self-consciousness etc., are peculiar to man, it may well be that these are the incidental results of other highly-advanced intellectual faculties; and these again are mainly the result of the continued use of a highly-developed language.’ (Darwin 1871: 103) The same applies to his contemporary Mueller (1887), who argued that the new ‘science of language’ dispenses with the need of a separate ‘science of thought’.
merely representing the study of a neurocognitive domain taken to be specific (*a* cognitive science among others).

## 2 The cognitive function of grammatical organization

### 2.1 What powers the generativity of thought?

Human thought, apart from expressibility in language, is ceaseless and has unbounded generativity. It essentially never stops in conscious human mental life, though not all of our mental activities have a linguistic format and this format can have various degrees of explicitness and articulation. Whether the brain is actively engaged in attending to a conversation or in solving some problem, or we simply let our minds wander, thoughts keep coming into our heads from we don’t know where, an ignorance dignified with the term ‘creativity’. Based on a rich storehouse of lexicalized conceptual knowledge (semantic memory capturing our general and multi-modal knowledge of trees, cars, people, hunting, cities, etc.: Binder et al. 2009), this thinking system generates ever-new thoughts flexibly, adaptively and expressed in speech. It thus has to be powered by a neural engine that provides it with the structures it needs and unifies them into thought-sized units. If this engine is separate from the linguistic one that we need independently, it follows that the conceptual resources on which it is based will not be lexemes (lexicalized concepts), and the principle by which such resources are systematically accessed and then combined will not be grammar. Driven to this extreme, the view entails that grammar may make no difference to what kind of meaning can be grasped by a human mind and conveyed in communication.

In this case an ape could in principle (ecological differences aside) think like us, just not say so – be it because he lacks the social motivations to do so, on which language is taken to depend (Tomasello 2008), or because relevant cortical-laryngeal connections enabling vocal control in humans (Deacon 2006; Fitch 2010) are contingently missing, leaving the thought system unaffected. In the latter case, its (his?) predicament would in some ways resemble that of post-stroke patients with brainstem lesions leading to ‘locked-in syndrome’. Like the hypothetical ape, such patients cannot operate speech or sign, but they can have minds like ours, as eye-controlled, computer-based communication technology can in these cases reveal (Laureys et al. 2005). The analogy breaks down immediately, however, when we realize that locked-in patients are speech-less (non-vocal) but not language-less (non-verbal). With eye-controlled
communication technology in place, a mind is revealed that can have a normal linguistic articulation internally. The hypothetical ape’s case is thus different, as he is meant to have Cartesian thought, whereas the locked-in patients have normal verbal minds. For an example of truly non-verbal minds we need to turn to approximately 25% of children at the low-functioning end of the autism spectrum, who fail to develop language in either production or comprehension in any modality, in the absence of physical impairment preventing language learning or articulation, such as deafness or apraxia (Norrelgen et al. 2015; Bal et al. 2016; Tager-Flusberg and Kasari 2013). These children, most of whom also exhibit intellectual disability (verbal or nonverbal) in the severe range, and are unlikely to develop language later, do not have minds like ours: they clearly do think differently, often failing to reach a representational (rather than sensorial) level of cognition altogether (Maljaars et al. 2012).

The hypothetical ape’s Cartesian thought system now becomes surprising. By hypothesis he is dissimilar from the locked-in patient, who has normal linguistically articulated thought, but he is dissimilar from a nonverbal child as well, since his Cartesian thought system is meant to be like ours, while a nonverbal child who has not developed language by school age has a profoundly different cognition. A more likely possibility seems to be that real apes are, unlike locked-in patients, language-less as much as speech-less. They lack the internally structured and articulable thought system itself, which goes with speech or sign and their grammatical structure in humans. In other words, where speech or sign are present, this is not merely contingently so (i.e., a fortunate accident), but speech/sign is foundationally significant for cognition. This would not deprive non-linguistic primates of cognitive activity, intelligence, or rich mental life, as there clearly are forms of perception, affect, categorization, navigation, imagery, reasoning, associative learning, and ‘mental representation’ (Gallistel 2009), none of which depend on language.

On this un-Cartesian view, speech (or sign) defines a species-specific cognitive type – it is the configurator of the sapiens-specific mind. Apart from the ‘language connectome’ (Vassal et al. 2016) in the brain – the set of cortical and subcortical regions involved in language processing together with their connections – there is no independent ‘mental’ process of complexity unfolding in it that would correspond to sapiens-specific ‘thought’ – though naturally, as noted, there can be other modes of cognition that do not involve thought in anything like the same sense, such as musical cognition (which lacks referential meaning and content, a defining feature of thought), arithmetic, navigation, or emotion (Fedorenko and Varley 2016). The view entails the predictions that in development and evolution, language will mature along with thought, implicating and shaping other cognitive systems such as working memory and forms of
executive control allowing flexible and rational behaviour; and that the loss of the relevant mode of thought will be the loss or disintegration of language.

No unidirectional causal arrow leading from language to thought is here assumed, which would make no sense on this view since language itself requires the form of thought in question and does not ever occur without it (except in pathological cases, which are of interest precisely for this reason). Nor does it entail some super-Whorfian form of linguistic ‘relativity’, which cannot even be articulated on this view, which maintains that there is no independent thought system, which could then in some respects be ‘relative’ to the specificities of a local language.\(^4\) While the un-Cartesian linguist targets the existence of this single thinking/language (‘languaging’) system as an *explanandum*, the Neo-Whorfian aim is to detect co-variations of this system (identified under the label of ‘thought’) with specificities of the local language. Any such specificity challenges the un-Cartesian account, insofar as we assume that the human cognitive type does not vary across the species and linguistic variation should therefore not affect it. However, the basic un-Cartesian suggestion here has been that the variation in question centers on lexical and morpho-syntactic dimensions of language, not on the forms of meaning that grammar is hypothesized to mediate on the un-Cartesian view (Hinzen and Sheehan 2013: Chapter 5). The claim, for example, is that a range of forms of object and temporal reference are grammatically configured in every language – but not that a given language has to use articles such as English *the* to implement forms of definite reference; nor that such forms require the existence of a morpho-syntactic category ‘Noun’; nor that the language would have to feature a morphological future tense. Neither does it affect the fundamental grammatical relations between arguments and predicates if Agreement shows up morpho-syntactically on the verb with respect to none, one, two, or more of its arguments.

The un-Cartesian hypothesis would explain why we never empirically find language developing without a particular mode of thought developing

---

\(^4\) Wolff and Holmes (2011) rule out what amounts to a variety of this view from the outset, based on four reasons: (i) people can have thoughts that are difficult to express; (ii) we can understand linguistic expressions that are ambiguous; (iii) words expressing new concepts could never be coined because their meanings could never be imagined; (iv) infants and nonhuman primates can think. Note regarding (i), that in fact we can only identify the thoughts in question after we have found a linguistic articulation for them; regarding (ii), that whether the expressions are lexically or structurally ambiguous, they are in fact disambiguated at the relevant level of linguistic representation; regarding (iii), that imagination is not merely powered lexically but also grammatically (compositionally); regarding (iv), that the relevant forms of thought, if indeed they were to qualify as such under more explicit and rigorous definitions of ‘thought’, are different.
alongside. Even conceptually, though, it is not clear how we could find the one system without the other. Language without the relevant kind of thought expressed in it would be a parody and hence could not be language in the same sense; and thought that was not articulable in a language in some modality would not be thought of the same kind. What could it even mean, therefore, that there is ‘thought without language’, apart from the following three facts, which I take to be uncontroversial here:

(i) Different types of what we might want to call ‘thought’ exist in different species as well as in ours,
(ii) In humans, language does not develop in an instant but unfolds in development over a number of years,
(iii) When we think, we do not necessarily speak or understand speech, as the cases of young infants, locked-in, deaf, or aphasic patients illustrate?

A widely acknowledged and inherent empirical feature of the relevant kind of thought, its intensionality, takes us beyond this conceptual argument, and back to the basic question raised above, of what powers the generativity of thought in the brain, if that is not the language network.

2.2 The intensionality of thought as evidence of its linguistic nature

All thought inherently has a content – without it, it would be empty and not be thought. This is one sense in which thought is intentional (with a ‘t’): it is always ‘about’ some object, person, state, event, possibility, or fact, and properties that these involve. Over and above such intensionality, however, the exact constituents and relations involved in a thought co-determine its identity. For example, if I say to you, ‘This man is my daughter’s primary school teacher’, then I am referring to a particular man and the thought that I here express is clearly different from the thought ‘This man is my wife’s secret lover’. If you told someone that I had expressed the latter thought when I had in fact only expressed the first, you may clearly be wrong. If you are, I would have referred to this particular person only under one of these two descriptions. Both of these thus may refer to the same person, but I may not know that, and even if I do, for the one description to figure in my thought is not for the other to do so as well. This is intensionality (with an ‘s’), a consequence of the fact that, in humans, reference is mediated by descriptions: it involves lexicalized concepts providing descriptive properties of an object, which may also non-transparently apply to the referent (the person in question is both my daughter’s primary school teacher and my wife’s secret lover) (see Box 1).
Whether or not a description applies is thus not a function of what you or I know or believe, or what descriptions or concepts we choose to apply – it rather depends on the world, in this case on who is the lover of whom. Intensionality requires a complex constellation in which meaning, belief, and world come together in a specific way.\(^5\) It can only exist when thought is indeed referential and when it is also conceptual/descriptive, i.e. when there are two ingredients that come together in a single referential act, which as such has both an external (i.e., the referent) and an internal (the concept/description) component. Having thoughts with that feature is an inherent aspect of any act of reference involving lexical concepts and hence virtually all thought.\(^6\) As I will argue in Section 2.6, these units of referentiality necessarily are grammatical phrases, i.e. they are linguistically complex in the way of a determiner phrase such as ‘the man’ in English, which combines a lexical concept (‘man’) with a grammatical functor regulating reference (‘the’). ‘Man’ as such cannot, as a lexical item, refer to a particular person, or distinguish between a particular man, the man I met, any man, or mankind – it is not as such referential, capturing a general lexical concept only.\(^7\)

Intensionality entails that when we change the lexical or grammatical (i.e. relational) ingredients of a thought, the thought becomes a different one, whether or not we keep reference (the external element) stable. This in turn entails that any generative system that was not language but generated the same thoughts would have to generate exactly the same lexical concepts and relations that we see expressed in normal human speech. Any change in these would correspond to a different thought. In other words, for each and every of the lexical or grammatical (relational) ingredients above, ‘mental’ (non-linguistic) equivalents would have to be found. Positing these should then not be circular

\(^5\) As Davidson (1982, 1997) classically argues, intensionality is a sign of thought as such, insofar as it indicates that reality is conceptualized and the concepts in question stand ‘in between’ the person referring and the object referred to. The contents of the thoughts are not determined by the objects, and have a specific identity determined by both the object and the concept nonetheless.

\(^6\) Qualifications concern the cases of pronouns, which involve no lexical concepts at all, and of proper names such as John when they have ‘rigid’ reference (reference relatively independent of any descriptive information provided, apart from the fact that the person referenced is taken to be named John). See further below and Sheehan and Hinzen (2011), and Martin and Hinzen (2014).

\(^7\) It is not, as such, a predicate either, as ‘predicate’ is a grammatical notion and this lexical concept does not need to function predicatively in grammar (as seen in kind-referential uses like Man comes from Africa).
or redundant, if a language-thought dichotomy is to be defended. What follows is an argument that no such alternative system exists.

To illustrate the circularity problem, consider a thought involving the concept SMILE (as opposed to any other concept) as an aspect of its identity as a thought. SMILE as such does not allow us to determine which or whose smile is intended, or even whether it took place; nor whether this concept will grammatically function in language (or thought) as a noun or verb. This rather concerns how this concept will function as a part of speech (PoS) on an occasion of its retrieval, which entails a crucial meaning difference. Thus, if I say (or think) ‘Mary smiles’, then what I say depicts an event placed in time and it expresses a proposition that is true or false; whereas if I merely utter ‘Mary’s smile’, then despite the absence of a substantive lexical difference, the grammatical difference between these two constructions entails that in the second case no proposition is expressed and instead only a certain smile is being referred to as if to an object. This formal ontology (i.e. event/proposition vs. object) is thus fixed grammatically, not lexically, showing that grammar is meaningful, though in a different way than the substantive lexicon, and also that it aligns perfectly with the formal ontology involved.\(^8\)

From a Cartesian point of view, one could seek to avert this conclusion by capturing the relevant difference through a notion of an ‘event’-concept vs. ‘object’-concept viewed non-linguistically. If such a cognitive distinction would be available to our minds – though note that referentiality in general is not a perceptual notion – the grammatical difference in question would not be responsible for the meaning difference. But this will only be coherent as long as the occurrence of each of these ‘event’ vs. ‘object’-concepts can be identified independently of the relevant grammatical distinction. Put differently, the difference between the mental ‘event-concept’ SMILE and the ‘object-concept’ SMILE needs to be some other difference than the one that we empirically see between this concept when it functions as a verb and when it functions as a noun. Surely it is circular to say, with Pierre d’Ailly in the fourteenth century, that the mental equivalent of a ‘noun’ is a ‘noun-like concept’ (Covington 2009: 125). It is also clear that there does not need to be any perceptual or external difference when I use the one construction over the other: clearly, both could be used in response to exactly the same perceptual input (a visual scene with a smile) (see Box 2). Grammatical meaning in the above sense

\(^8\) This is often denied in the PoS literature, where the common view is that while there is a semantic basis for parts of speech distinctions, namely the object-event distinction assumed to be given non-linguistically, this basis does not suffice and there is an additional morphosyntactic dimension to the N vs. V distinction that cannot be grounded semantically. See further Hinzen and Sheehan (2013: Chapter 2).
therefore does not co-vary with either perception or denotation, neither of which can as such determine which construction will be used.\textsuperscript{9} This challenge, illustrated here for the simple Noun-Verb distinction, multiplies with any further element of grammatical structuring that influences what thought we are thinking, beyond the lexical material that it contains (ignoring other information that can matter to the identity of a thought as well, such as phonology, prosody or affect).

In sum, the challenge to Cartesian thinking is that any element of a given thought can matter to its identity, but the only system where we see these elements exhibited in the right configuration is the linguistic structure that corresponds uniquely to the thought in question. The suggestion is that human grammar effectively is the single known system that integrates reference and description in the right way, in units of structure (phrases) that necessarily involve them both. The elements in question therefore either are elements of language as it is processed in our minds and underlies our actions, or else they must trivially replicate them at a separately posited ‘mental’ level. If so, the null hypothesis should be that grammatical organization in humans is (rather than merely expresses) the thought system that exhibits the form of referentiality in question, with no independent generative system required.

This in turn illustrates in just what way grammatical organization through phrases and sentences is inherently meaningful, subserving a species-specific referential function and thus contributing a kind of meaning unavailable lexically or non-linguistically. The core cognitive function of grammar is to turn lexicalized concepts, which as such do not refer (but are categories of perceptual experience encoding general concepts encoding semantic memory), into referential expressions (of different types) on an occasion of language use. We may call the result grammatical cognition, insofar as it reflects a species-specific cognitive type from which grammatical organization cannot be subtracted without destroying the cognitive type itself. Note that grammatical cognition by its very nature is ‘social cognition’, not merely because reference is involved but also because it requires words which are learned through sensory-motor integration between information that we hear (from others) and then express. This mode of cognition is thus inseparable from a linguistically re-structured social space in which such learning can take place. This, in short, is the un-Cartesian answer to the question of the cognitive significance of grammar. Below we will review evidence that:

\textsuperscript{9} As the Modists put it, not the \textit{significatum} differs, but the \textit{modus significandi}.
(i) The same form of referential meaning is not known to be available to non-grammatical species, and it matures in humans along with grammar.

(ii) It is available to all neurotypical humans and mediated grammatically in all languages in systematic ways.

(iii) It is systematically disturbed in mentally disordered populations, where an associated disturbance in the thought system and social cognition are also seen.

2.3 The demise of ‘concepts’

If SMILE is a concept, then ‘object’ or ‘event’ in the sense of the above discussion are not further such concepts. Rather they correspond to two possible ways in which given lexical concepts can be made referential on an occasion of their use, depending on the grammar that is put in place. This ‘grammaticalization of meaning’, which results in forms of reference and a formal ontology of thought, therefore does not add further concepts. Grammar is not a mapping from conceptual space into conceptual space, but from given concepts to something novel, namely reference. Yet, what kind of semantics is there prior to the point in evolution or development where concepts have come to occur as lexical items in grammatical structures? What about meaning at the ‘root’ level, before there even are PoS distinctions (Harley 2014)? Why could the meaning of a word not be a pre-linguistic abstract ‘concept’, existing independently of language and its grammaticalizations? Or do we only see such concepts where we also see them used in grammar?

Such pre-linguistic concepts would be non-lexemic by definition as well as factually, since non-linguistic species do not lexicalize the categories they form based on their experience, in the sense of creating freely available and shared sound/form-meaning pairs that can be activated independently of sensory experience for thinking and talking purposes. Nor would such concepts be used referentially, on the assumption that evidence for (declarative) reference in a human-specific sense in non-linguistic species or pre-linguistic humans is extremely limited (see further Section 2.5). The non-linguistic concepts in question would thus function very differently from lexemic concepts in humans, lacking two of their hallmarks, which raises the question why the same technical term should be used for both.

Our primary evidence for a productive form of concept combinatorics, moreover, comes from the case where these concepts are grammaticalized, as in a structure like (i):
(1) brown cow

Here what we see are two lexemes, one in the role of a modifying adjective, hence denoting a property, another in the role of head noun, determining the referent of the phrase as a whole. This is why the meaning is what it is: a cow that is brown, not a cow-type of brown, which is what would be denoted by (2):

(2) cow brown

It could be argued, then, that the alleged ‘concept combinatorics’ illustrated in (1) is in fact a word combinatorics and requires grammar. Evidence is required for positing the same combinatorial process in the absence of the lexemes that we here see, their grammaticalization, and their referential use. What would be the advantage, then, in humans, of positing such a ‘mental’ combinatorics, when the lexemic one suffices, and if on the ‘mental’ level the exact same distinctions would have to be re-stated? No matter how open we are towards the idea that non-human species can think, it is (i) not lexemes they combine, and (ii) there is little or no evidence for either grammar or compositionality. Even in humans, concepts in the sense of lexemes do not in fact themselves combine with others – i.e. directly. Rather we see them appearing as PoS, which (outside of the NP-internal structures depicted in (1)-(2) above) are in turn first combined with a grammatical ‘edge’ created by a functional morpheme, as in (3), unlike in (4), where no such functional material intervenes in lexical combination:

(3) [The man] [walks [a dog]]

(4) [man [walk [dog]]]

Suzuki et al. (2016) have recently claimed the ‘first unambiguous experimental evidence for compositional syntax in a non-human vocal system’ in Japanese great tits (Parus minor). But the evidence concerns composite behaviour, not compositional meaning, and involves no evidence for headedness, an inherent aspect of any word combination in humans. As for Arnold and Zuberbühler’s (2008) evidence for ‘meaningful combinatorial signals’ in primates (Putty-nosed monkeys), it involves no compositionality insofar as the meaning of individual calls is not preserved, as the authors note. Schlenker et al. (2016) show that the formal tools of modern semantics allow giving an alternative compositional semantics of the same data. If so, these tools cannot distinguish between monkey and human meaning, the latter intentional, intensional and declarative inherently, the former inherently action-related and premised by specific constellations involving threats and arousal. My thanks to Joana Rossello for discussions on this issue.
Grammar thus appears as the specific glue between human concepts that makes them combine productively.

As stressed earlier, none of this means that there is no meaning pre-linguistically. In particular, content words in the human lexicon such as MAN, EDIBLE, WARM, etc. can be used to classify perceptual data as falling into abstract stimulus-classes. In this sense they correspond to perceptual categories. Members of these categories statistically exhibit certain perceptual features (e.g. HAIRY, BIPEDAL in the case of MAN), and they form hierarchical relations, giving rise to a semantic field or network. I will here assume that these categories are either associatively learned from patterns of stimuli occurrences and rewards, or they are innate categories, including foundational abstractions such as ANIMATE, NUMBER, AGENT, OBJECT, INTENTION, CAUSE, etc. (Gallistel 2009). The latter are rich and profoundly abstract in content, in part relational (e.g. CAUSE), not specific to humans, and they structure the experienced world before reward-driven learning starts, laying the foundation for a particular type of experience as such. But they differ from lexemes (like the words seven, intend, or cause), which are not stimulus-determined in the way that percepts are and structure our thinking independently of what we perceive at the moments when these thoughts occur. Although words are given to us and we are bound to use them with the meanings they have, the sentences containing them are something that we choose to configure on a given occasion, and we are accountable for this, unlike in the case of percepts. Any object that we think about, we can also think about under different lexemes, or not at all. Sentences can be false, and we use them being sensitive to that. False perceptions exist, but differ from false assertions.

In short, non-human species have brains that can compute, represent, and perceptually classify, but lexical concepts are no more required for this than reference is. Neither is speech (or sign), on which lexemes depend in humans. Even if lexemes existed in a non-human species, this would only give us content of one particular kind: the kind of content through which stimuli fall into particular classes. Moreover, combining these with a generic and formal operation such as the Merge operation of current Minimalism (Chomsky 1995, 2007) would not as such contribute any additional meaning. So Merge as such plus the lexical content in question will not give us the distinctive kind of content that sentences as opposed to words carry (grammatical meaning), unless we were to add an additional and independent principle. We could add semantic

11 Nor that at one stage of language evolution, a ‘proto-syntax’ might not have existed dispensing with the kind of ‘shell’-structure seen in (3), which embeds phrases in functional layers (see Jackendoff 2002; Progovac 2015; on ‘syntactic fossils’).
compositionality in particular, defined so as to yield propositional meaning from lexical content of the MAN, EDIBLE, WARM type. But this would presuppose compositionality, since the compositional mapping depends on a function defined so as to yield the propositions in question from lexemic input (Davidson 2005). Designating some lexical items (i.e. verbs) as functions and defining them as mappings from objects to the propositional meanings in question is circular. No ontology of objects of reference is even expected to arise in this fashion, as lexemes only have general meaning and even compounding, as in HOUSE BOAT, only ever yields another general concept. In short, grammar as an independent principle behind human-specific forms of meaning is required. If we helped ourselves to another type of lexical item, namely function words such as the, and define a suitable compositional meaning for these as well, then such items already express grammatical functions – they do not correspond to perceptual classifications.

In sum, grammar and the kind of meaning it encodes cannot come from the lexicon or from pre-linguistic cognition, if these systems lack lexicalization; but also not from lexicalization, since lexicalization without grammar does not add up to referential and propositional meaning. If these pre-linguistic systems did contribute this different type of meaning, on the other hand, then they would already master something that we de facto only see occurring with language (and grammar). The material cultural records of non-sapiens hominins (Tattersall 2008) however do not suggest that these species could think the same thoughts and had concepts as we express in language. ‘Thought’ is a species-specific category. It varies as species do (Penn et al. 2008), even within the genus Homo, and more certainly beyond it.

If a new type of content or thought arose with language, pre-linguistic thought systems would not be able to ‘read’ this kind of content, just as vision cannot read propositional information, though some top-down influences also exist. If an interface between pre-linguistic thought and language (Chomsky 1995) is posited, therefore, the prediction would be that structures built linguistically would arrive at the semantic interface, which the systems on the other side of this interface could not use – it would be like a strange and foreign code. To the extent that the intrusion of language into the hominin brain changes what thoughts it can think, language design cannot then to that extent be rationalized as a way of meeting ‘conditions’ imposed externally at some interface by extra-linguistic semantic systems (Chomsky 1995). Hence either the grammaticalization of meaning (in the present sense, i.e. the fact that meaning becomes structured grammatically) makes no difference to what meaning or content is available to a mind, in which case notions of interfacing thought systems plus a notion of a meaningless syntax will suffice. But then we have no
explanation for how propositional meaning arises, which non-linguistic species
do not possess, if we understand that to be the kind of meaning expressed in
sentences. Or grammaticalization does make such a difference. Then there is a
kind of meaning that only arises with and within grammatical structures –
grammar is transformative for what kinds of meaning there are. In that case
there will of course be interfaces with non-linguistic perception, learning, and
affect, but extra-linguistic thought systems will neither externally condition nor
‘read’ the deliverances of this new type of cognition.

An independent mental world of concepts in humans that are not lexemes
does not only seem unnecessary and redundant, but it is unclear what its neural
correlates could be. Binder et al. (2009) set out to identify the ‘semantic system'
in the brain that underlies our understanding of word meaning. Results were
based on a meta-analysis of 125 studies investigating verbal semantic processing
based on a stimulus presentation in meaningful words as opposed to externally
and stimulus-driven processing of percepts deprived of conceptual, word-based
meaning (e.g. pronounceable pseudo-words). Results indicated an extensive and
widely distributed network in the brain minimally overlapping with the com-
plementary perceptual/external one. Together, these two make up much of the
human cortex. Where would non-verbal mental concepts be, equivalent to word
meanings yet different from them in being language-independent?

Turken and Dronkers (2011), comparing Binder’s network to their own
auditory ‘language comprehension network’ based on a functional connectivity
profile of the left posterior middle temporal region, find both networks ‘largely
consistent’. Binder et al. (2009) further consider their network ‘strikingly similar’
to the default mode network in the conscious resting state (Raichle et al. 2001),
suggesting that the brain’s default mode (prior to an interruption through
external task demands) could itself be highly linguistic in nature, involving
the ongoing retrieval of conceptual knowledge. Both this default network and
Ferstl et al.’s (2008) ‘language comprehension network’ in turn comprise classi-
cal ‘theory of mind’ regions. Neuroanatomical and functional investigations of
language in the brain have moved beyond the localizationist flavor of early
neurological models based on lesion studies, revealing a widely distributed
neuronal network held together by a complex set of white matter tracts con-
necting structurally and functionally remote brain regions as required for sen-
sory-motor and conceptual integration (Price 2000; Vassal et al. 2016; Duffau
et al. 2014). Language processing also implicates both hemispheres (Mohr et al.
1994) and cortical-subcortical loops critical to its functioning (Lieberman 2007).
Why posit a network of non-linguistic concepts or thoughts, if a sensory-motor
system implicating grammar and a conceptual-lexemic one (semantic memory)
sufficed for the functional thought system that we see in humans?
2.4 Language as triangulation

This, of course, depends on our view of what language is and encompasses. Only then can we tell whether an extra system is required. To start with, language is an integrative system. In every utterance we ever make, multiple neurocognitive domains are involved and integrated in a unified and coherent fashion, such as memory, executive function, perception, attention, meaning and affect. Also involved is selfhood, since people producing utterances act as persons not robots, and more specifically as first-person referents for themselves. Even where the grammatical 1st Person is not explicit in overt speech, utterances entail the knowledge ‘I say/think that…’, which is to say that speech contents and acts are subordinated to a self as identified in the 1st Person. Persons who experience what they say as ‘he says/thinks that…’ would have a disturbance of the self, corresponding to a symptom of schizophrenia: the delusion that one’s thinking is carried out by someone else (Schneider 1959; Crow 2010; see further Section 3.2). Acts of reference marking a given referent grammatically as 3rd or non-Person (‘he’, ‘the man’, ‘it’, ‘the world’), too, make reference to Person implicitly, as any such act is enacted by a 1st Person for a 2nd, with 3rd Person singled out as neither 1st nor 2nd. Moreover, who is being referenced by ‘John’ or ‘the table’ depends on who is uttering it and identifying himself in the 1st Person while doing so. When Bill said ‘John left the table’, he may be taken to have referred to a person known to him and the interlocutor as ‘John’, as well as to an event of leaving a particular proximal table known to both, anchored in time as something that took place and was completed prior to the point of speech. As a result of that, the utterance is true or false, and true only if the relevant person and event both exist and stand in the grammatical relations to the 1st Person and the hearer that are grammatically specified.

Person specifications thus act as a system with all three grammatical Persons co-present in the form of a ‘triangulation’ as depicted in Figure 1: A 1st-Person self, connecting to a 2nd Person via the baseline of speech, puts forward a thought in language about the world viewed as independent of both speech agents.12 Note that the meaning of ‘1st Person’ is crucially not that of ‘the speaker’, in the sense that ‘I say...’ and ‘The speaker says...’ have different meanings, even if I am the speaker. In the second case, ‘The speaker’ can refer to me, but it need not (any speaker could be referenced), while in the first case, it must. Thought taking place outside of this triangulation, if it deserves the name, is different in nature, and can have different baselines.

12 See also Davidson (1982, 1997), who on the other hand does not relate this triangulation to distinctions of grammatical Person, which are crucial here.
That reference is grammaticalized in humans as stated earlier can now be seen to follow from the necessary involvement of grammatical Person specifications in acts of reference alone. Note further that grammatical Tense has no meaning outside of the actual speech events in which it is specified, anymore than Person does. An inherent aspect of the triangulation, therefore, is speech. Interpreting Tense requires witnessing the speech act, as and when it takes place, since grammatical Tense encodes a relation between the event that the speech act is about and the time of the act. Since specifying values for both Person and Tense is a grammatical requirement on the well-formedness of any sentence, and they both interact with core dimensions of grammar such as structural Case, grammar is inherently integrated with speech. Without fixating temporal reference and Person, no thought content can be fully determined either. Thought content, too, therefore, depends on speech, which is not an accidental and fortuitous extra appended to an otherwise already functional thought system.

### 2.5 The unfolding of the triangulation

In neurotypical humans, communicative acts exhibiting reference in this triangular sense first shows up cross-culturally when infants start to spontaneously produce declarative pointing gestures from around 10 months of age (Cartmill et al. 2012; Butterworth 2003), such as an index-finger point to an object accompanied by exchange of eye gaze with a 2nd Person whose attention the infant draws to the object and shares. Such intentional acts of gestural communication

---

**Figure 1:** The deictic frame of speech and thought. A grammatical first Person, addressing a second, talks about an object or event in the world (the ‘it’ or third/non-Person) expressing a thought.

---
are often characterized as non- or pre-verbal (Tomasello 2008; Csibra and Gergely 2009). However, infants starting to behave in this fashion are not only genetically linguistic but, in comprehension, infants’ vocabulary is well underway by this time. Even in production, their gestures are often accompanied by either babbling or the first one-word utterances, e.g. ‘dog’ when pointing to the respective animal, and their gestures are prone to elicit verbal responses from their caretakers. The size of children’s vocabulary and the number of different kinds of objects pointed to at the babbling stage moreover correlate (Iverson and Goldin-Meadow 2005).

The co-presence of words in declarative gestures suggests that the objects that the child points at are identified as being of a particular kind, i.e. as falling under a certain description or general concept as captured by the word uttered. In fact, infants point more when the adult’s response is informative about the object, rather than merely consisting in sharing attention to it (Kovács et al. 2014). The concept of joint attention (JA) as such does not capture this predicational dimension of the pointing act, which appears more related to language. In comprehension, too, there is evidence that one year olds understand declarative referential acts involving deictic words together with a symbolic description, and to be appreciated in this way, the deictic and the descriptive elements not only have to occur concurrently but also to originate from the same source (person), in a single integrated referential act (Gliga and Csibra 2009). In sum, general concepts lexicalized as words are in place when reference begins, and they are predicatively combined with the declarative gestures (Cartmill et al. 2014; Cartmill et al. 2012; Goldin-Meadow and Butcher 2003), forming integrated units involving both reference and description in different modalities.13

Studies have also revealed longitudinal correlations between different types of declarative co-speech gestures and later forms of grammatical complexity in infant speech. Children first produce words combined with pointing gestures that ‘reinforce’ the gesture – for example, pointing at a dog and saying *dog* – and only later produce words that ‘supplement’ the information contained in the gesture – for example, pointing at a dog and saying *eat*. The individual onset of reinforced co-speech-gestures predicts the individual onset of determiner phrases in speech, while supplemented co-speech-gestures predict the individual onset of sentences in speech (Iverson and Goldin-Meadow 2005; Cartmill et al. 2014; Özçalışkan and Goldin-Meadow 2009; Goldin-Meadow and Butcher 2003). As noted, verbally ‘reinforced’ gestures already exhibit the format of a

---

13 Reference remains bimodal even in its fully grammaticalized form: noun phrases like ‘that man’ cannot be felicitously produced in their non-anaphoric uses without an accompanying co-speech gesture.
proto-grammatical configuration, since we can clearly distinguish the pointing gesture itself, as the referential part of the communicative act, from the word uttered, the predicative part (the concept applied to the object referred to). This is the same duality that we later see mono-modally in determiner phrases (DPs), where we see an ‘edge’ regulating the reference and an ‘interior’ where a lexical-descriptive content is located:

\[
\text{DP a/the [NP DOG]} \quad \text{edge} \quad \text{interior}
\]

In short, a language-specific triangulation is falling into place by the time reference starts. In fact, its earliest manifestations are seen long before the first referential gestures are produced or understood. Auditory learning starts in utero, and sensitivity to phonological properties of the ambient language is not only reflected in the newborn’s perceptions and preferences, but their very first natural vocalizations: even crying is an aspect of language development (Mampe et al. 2009). Newborns also exhibit a preference for listening to speech as compared with complex non-speech analogues that control for critical spectral and temporal parameters of speech (Vouloumanos and Werker 2007). Attention is preferentially ‘tuned’ to language, ensuring that from the beginning language is there to structure the infant’s learning and social interaction, drawing its attention to the communicator long before its visual capacities mature.

It is not merely that very young infants like speech more than other complex sounds, but that from very early on they are sensitive to intrinsic properties of the kind of system that language is and the meaning it carries. 6-month old infants, even when they do not understand anything that is said, have an abstract understanding of the fact that in such speech (but not non-speech, e.g. coughing), information is communicated about an object, and hence involves reference (Vouloumanos et al. 2014). In another study, 4-month olds were demonstrated to have referential expectations for speech but not backwards speech, even when they fail to understand the conceptual content of either of these (Marno et al. 2015). In this way the deictic space involving a baseline of language connecting two persons and a referential content is being set up from the day humans are born. There is no time when humans only ‘think’ and then eventually they also have language; or when they have ‘social cognition’ but no ‘linguistic cognition’.

Finally, reference as first manifested in (declarative) pointing and as correlated with grammar is also absent in non-grammatical beings. This has been widely argued to be true of chimpanzees, who understand gaze and head
movement, following an experimenter’s line of sight even when it projects outside their perceptual field, but do not point declaratively or understand declarative pointing as human infants spontaneously do (Butterworth 2003; Tomasello 2008; Tempelmann et al. 2013). Animal communication in the wild can exhibit forms of functional reference, but as noted with crucial differences to human reference. An animal alarm call does not seem to involve meaning in the sense of a predicative concept providing generalizable information about the referent as an instance of a kind: it has a functional meaning consisting in a course of action to be taken in a predatory incident witnessed by the caller. Moreover, it is not lexicalized in the same way as words are, which combine with others and can be retrieved independently of stereotyped stimulus contexts causally triggering their use (Fitch 2005: 205, 212; Bickerton 2009: 44–47, 68–69; Cheney and Seyfarth 1990; Deacon 2006). While communicative gestures in the great apes are more clearly intentional and flexible, they are neither combinatorial nor triadic, intensional, or discrete, regulating social interaction but failing to convey shared referential and predicative content of the type that even first words in humans exhibit (Tomasello 2008).

2.6 The grammaticalization of reference in humans

In the ‘edge’ of the referential template in (5), material without substantive lexical content is added to the NP. This material functions so as to regulate the way in which the content expressed by the NP is used to refer: the phrase ‘the dog’ can (occurring as part of an utterance) function referentially in a way that DOG as such, i.e. as a mere lexeme or general concept or category, cannot. Without creation of an edge, then, lexemes cannot function referentially or implement referential distinctions: non-grammaticalized, DOG as such cannot distinguish between a dog I saw, this dog, the dog I met, some dogs, dogs in general, dogs that I like, dog as mass (as in ‘I ate dog’), or dog as a species. Reference in this sense, although it depends on a lexicon, thus falls on the grammatical side and is technically a grammatical concept, which we cannot locate in any non-linguistic ‘conceptual-intentional systems’. As noted, languages differ in the morpho-lexical resources with which they implement such referential distinctions. But the more the possible forms of reference are indeed grammatical in nature, the less we would expect that the same lexical resources need to be available in order for the same grammatical functions to be carried out.

Grammar is not only required for reference in the present sense, but grammar never seems to function so as to do anything else. Nothing but making a given lexical concept referential in some way can happen in the grammatical
process. This process has reached its outer limits when a truth-evaluable structure has arisen and a sentence is uttered functioning as a speech act with illocutionary force. In (6), the gradual ‘grammaticalization’ of the lexical concept *dog* is depicted:

(6) a. I had (some) *dog* (cf. *I had beef*)
   b. I like *dogs*
   c. I want to be a *dog*
   d. I saw (some) *dogs*
   e. I once had a *dog*
   f. I saw the *dog* (again)
   g. I like this *dog*
   h. I like him (*dog*)

In (6a) we see the lexical concept DOG becoming an argument with a possibly empty determiner/edge and the only possible interpretation ‘I had (some amount of) dog-meat’. Then, with plural Number specified, we see its denotation individualized, though reference is still generic, namely to all instances of the kind ‘dog’ (b). Then reference is to an arbitrary instance of a property (being a dog) (c); then to some particular instances referred to with an existential commitment (d); then to a specific such instance (e); an instance that has been mentioned before using the strong determiner ‘the’ (f); and finally, to a dog present in the speech context proximal to the speaker and pointed out with a deictic gesture (g, h). In each of these types of noun phrases, nothing other ever happens than that the same lexical concept is made to function referentially in one of a number of grammatically possible ways, with an increase in referential strength as we move from the most lexically mediated or descriptive form of reference to the most grammatically mediated or deictic forms.

Specifically, as the edge expands – from an edge that is necessarily lexically zero to one that has a referentially weak, indefinite determiner, to one with an obligatory ‘strong’ and definite determiner, to one with a deictic determiner or pronoun that *can* occur without the lexical restriction, and finally to a deictic pronoun that *must* occur without such a restriction – we see reference getting more specific, and stronger, i.e. less dependent on a lexical description or predicate.14 In the middle range of this hierarchy (i.e. c-e), the NP-restriction

14 Sheehan and Hinzen (2011) incorporate proper names into this hierarchy through an adaptation of the Longobardian idea that in rigid reference, the NP itself occupies the edge position (see Longobardi 1994, 2005). The weak/strong distinction along the hierarchy in (6) is related to
providing descriptive content is first still necessarily satisfied for the act of reference to succeed: If I utter (c), for example, then the use of the lexical description ‘dog’ is essential to what I say; if I utter (f), on the other hand, and whatever I refer to turns out not to be a dog, reference can still have succeeded (Donnellan 1966): it is only a presupposition that is violated, and the definite description can be used referentially without the description being satisfied.\textsuperscript{15} Demonstratives can already occur without a description altogether, and deictic pronouns must so occur. Reference thus becomes less lexical and more grammatical as we go along this hierarchy. Speaking of a hierarchy of reference is justified insofar as reference first is inherently predicative and cannot be individually specific, then can be specific but cannot yet be definite, then becomes definite and eventually deictic, until finally no predicative or descriptive component is left.

This account depicts grammar as a device of extended deixis, which expands the deictic possibilities of the index finger as used by one-year olds to the whole spectrum of humanly possible forms of object reference, beyond which there appear to be no other such forms. By creating a nominal edge and making it stronger/expanding it, a given lexical concept is embedded in deictic space. As we go beyond the nominal phrase, we first extend deictic embedding in space to embedding in time, which happens as verbs are added to nouns and objects become part of events with a temporal location specified relationally with respect to the time of speech through Tense. Finally, the events become part of propositions, which now can engender forms of reference that correspond to expressions being true or false. The essential divisions in a grammatical derivation, therefore, as depicted in (7), correspond to a set of transitions from one formal ontology of meaning to another, each of which inherently contains the previous one as a part (8):

\begin{align}
(7) & \quad \text{[Clauses [Tense/Verb phrase [Noun Phrase]]]} \\
(8) & \quad \text{[Proposition [Event [Object]]]} \\
\end{align}

Grammar in this way sets up a framework for thought. There is no thought of the human-specific kind without a formal ontology of this sort. What we refer to and think about is always specified for its formal ontological status: whether it is an

\textsuperscript{15} As Donnellan put it: ‘the speaker presupposes of some particular someone or something that he or it fits the description’ (Donnellan 1966: 288).
object, event, proposition, etc. One way of re-stating the un-Cartesian hypothesis is thus that there is no such formal ontology in thought other than the one we see arising empirically as grammatical complexity is built in acts of reference on an occasion of language use. Grammar not only provides for a framework in which thought can take place, but it also exhausts the space of possible thought, which does not have an ontology other than or additional to the one we see correlating with grammar. Since the limit of grammar is the configuration in which a truth value can be assigned, and beyond this there is no known additional formal ontology, the limit of grammar is also the contingent limit of thought.\footnote{16}

Fixing reference, then, is what the grammatical process converges towards, making use of predication; it is also the distinctive meaning that goes with grammatical organization, i.e. the content of grammar. But reference is not ultimately fixed in phrases, either. It also depends on grammatical relations.\footnote{17} Thus in (9), and (10), the exact same phrases are present in (a)-(d), respectively, yet their referentiality differs in each case, playing the role of sentential predicate in (a) and being necessarily low-scope in (b), while being referentially specific in (c) when the same phrase is the grammatical subject of an episodic matrix verb, and becoming quantificational when the same phrase is in the scope of a generic operator:

\begin{align*}
\text{(9)} & \quad \begin{array}{l}
\text{a. I am [a man]} \\
\text{b. Jenny would like to be [a man]} \\
\text{c. [a man] entered the room} \\
\text{d. [a man] always does the dishes}
\end{array} \\
\text{(10)} & \quad \begin{array}{l}
\text{a. I am [the man]} \\
\text{b. Jenny would like to be [the man]}^{18} \\
\text{c. [the man] I saw this morning had a hat} \\
\text{d. [the man] always does the dishes}
\end{array}
\end{align*}

Since the same referential possibilities transpire with ‘a man’ and ‘the man’ depending on the grammatical relations involved and these differences are systematic, the cause cannot be a lexical ambiguity in ‘a’ and ‘the’. Rather, it is grammar that contributes the meaning changes in question. Hinzen et al.\footnote{19}

\footnotetext[16]{This aspect of the un-Cartesian hypothesis is a strongly Wittgensteinian theme that is especially developed in Hinzen and Sheehan (2013: Chapter 9).}

\footnotetext[17]{These crucially include structural case as argued in Hinzen (2014b).}

\footnotetext[18]{Said, for example, by a casting agent distributing roles in a play.}
argue that the same conclusion follows for other core semantic phenomena such as intensionality, which is configurational as well, occurring when and only when a clause appears in the right grammatical relations. These alone determine whether a given clause is interpreted intensionally (with respect to a mental state) or extensionally (with regards to the world). In the same direction, Sheehan and Hinzen (2011) argue based on cross-linguistic facts that factivity, too, i.e. the interpretation of a clause as referring to a fact presupposed by the speaker to exist, has an inherent grammar, aligning with configurational rather than non-grammatical facts. On this model, the possible forms of reference available in the clausal case form a hierarchy as well, paralleling the nominal hierarchy in ranging from indefinite forms of reference (thought contents expressed by clauses subordinated under fully intensional verbs such as ‘believe’), to definites (fact-referring clauses creating reference presuppositions), and finally rigid ones (truth value-denoting matrix clauses).

If so, with grammar and the layers of complexity it inherently generates, a formal ontology of meaning falls into place that spans the space of possible thought and reference. It remains to specify one last instance of this, namely reference to objects that are also persons.

2.8 Language and person

The opening shot of modern philosophy was Descartes’ insightful classical dictum (11):

(11) Cogito ergo sum (I think, therefore I am)

Crucially, he could not have made the same point saying (12):

(12) Cogitat ergo est (He thinks, therefore he is)

In fact, he could not have said that even of God, since as Descartes reasoned, everything other than our own thinking is less certain for us than it, including God’s thinking. So the grammatical 1st Person is crucial, and in discussions of personhood in philosophy it has been ever since. Consciousness is often virtually identified with 1st Person thought (Chalmers 1996). The relevant concept of Person, however, is clearly here not the lexical one, which is related to a perceptual difference, in humans, between humans, who we perceive as persons, and (most) non-humans, who we do not so perceive (unless they are talking and personified, as in cartoon strips). The lexical concept PERSON is in
this respect like ANIMATE: a foundational abstraction structuring human experience. As such it has no relationship to the grammatical notion of 1st Person, which is the one relevant to the distinction invoked in discussions of personhood and consciousness since Descartes. The un-Cartesian proposal is that the use of the 1st Person in this second sense is not merely a contingent feature of external linguistic expression but essential to the phenomenon of consciousness and human-specific selfhood itself (Martin and Hinzen 2014; Hinzen and Schroeder 2015). In short, the grammatical distinction in question is a cognitive distinction, which in turn is critical to selfhood. Only objects that we refer to using grammatical Person distinctions are persons and exhibit normal selfhood in a neurotypical, sapiens-specific sense. It’s not that we first are persons, and then either language comes along or not. Rather, whether it does and how it is used matters to whether we are persons in the same sense. Note that both ASD and SZ precisely are characterized as selfhood disturbances (see Section 3).

This strong claim could be challenged on a number of grounds, for example by specifying the meaning of the word ‘I’ without circularly re-using the notion of 1st Person. To my knowledge, however, such attempts have not only not been made but a large literature centering on the notion of ‘essential indexicality’ (Perry 2000) indicates why such attempts are bound to fail. ‘I’, even when uttered by Tom himself, never means the same as ‘Tom’, ‘the speaker’, ‘this body’, etc., even if these are used by Tom himself. That none of these forms of reference can substitute for ‘I’ without a change in meaning was classically observed by Kaplan (1977), who looking into a mirror may notice (13) while (at least for a moment) failing to notice (14):

(13)  This guy’s pants are on fire
(14)  My pants are on fire

Crucially, there is self-reference here in both cases, which therefore is clearly not the point. The point is that we need self-reference in the right grammatical Person: the shift in Person from (13) to (14) indexes a cognitive distinction and an epistemological difference. This is how the point has not been analyzed, however, to my knowledge, due to the Cartesian view that the observation just made simply concerns the contingent way in which the relevant mode of self-reference happens to be expressed in language. Yet an analysis of the alleged non-linguistic ‘thought’ process that would not invoke the grammatical distinction in question would be hard to give. This result is expected given that ‘I’ invokes more grammatical complexity (specifically, deixis and the grammatical Person system) and less lexical descriptive content than any 3rd Person form of
It is therefore unexpected that 1st Person forms should be replaceable by the 3rd Person ones, without a change in meaning. ‘He’ in particular involves a specification of Gender and ‘this man’ uttered by myself specifies me as being a man. ‘This body’ merely specifies a physical object, reference to which does not equate to 1st Person thought, as the case of pronoun confusion in ASD illustrates (see below). There is thus no escape from the 1st Person when the relevant epistemological distinction is to be made. That reference to oneself as ‘I’ involves no description and is minimally lexical and maximally grammatical in this sense also explains the traditional epistemological observation that such reference is immune to error: There can be no error, when no description is involved that can be wrong.

There is a rationale for essential indexicality, then, which is grammar. Grammar is not only required for 3rd Person reference to objects, but for personal reference as well; and forms of the latter are ‘essential’ because there is a grammatical difference between personal and non-personal forms of reference, which correlates with a difference in the grammatical meaning and complexity involved.

### 2.9 Section summary

Acts of reference in humans depend on units of referentiality integrating a referential edge with a lexico-conceptual interior. As grammar builds complexity, these units become parts of others, with grammatical relations established between them that provide a **formal ontology** for thought. Core phenomena of semantics, namely referentiality, deixis, formal ontological distinctions, intensionality, and factivity, **co-vary** with grammar: they are **configurational**. The core cognitive function of grammar is the conversion of lexically contentful but referentially inert general lexical concepts into expressions that are referential in a number of grammatically possible ways, on an occasion of their use. A three-fold grammatical Person distinction involved in all acts of reference makes this referential space specifically triangular. Once a human-specific deictic space is spanned in this fashion, there is arguably no need, room or evidence for an independent thought system in humans. This hypothesis opens a research program which draws its motivation from a variety of sources:

---

19 As shown in Martin and Hinzen (2014), the morphological complexity of the Romance object clitics mirrors this increased grammatical complexity.
(i) Evidence for the configurational nature of reference in a human-specific sense, including all forms of reference from generic to personal;
(ii) The apparent absence of any of these forms of reference in species lacking grammar;
(iii) The fact that all concepts in humans (where they are lexicalized) function referentially (when they are grammaticalized) and none do in non-linguistic species (where they are not lexicalized);
(iv) The intensionality that the relevant type of thought exhibits in addition to its intentionality, and the absence of a system other than language that provides appropriate structures for contents of this kind.
(v) The developmental co-presence of a language-mediated form of thought from the day humans are born.
(vi) Disturbances of reference in disorders of thought under conditions of genetic variation within our species, to which we turn now.

3 Two disturbances in the human cognitive type

3.1 Autism Spectrum Conditions (ASD)

3.1.1 Symptoms of ASD and their relation to language

ASD is a neurodevelopmental disorder involving a fundamental change in how neurotypical individuals communicate, socially interact, and behave. Rather than leading to a generalized cognitive impairment or mental retardation, however, it involves a shift in ‘cognitive style’ or the balance between our normal cognitive functions. This shift can feature exceptionally enhanced lower-level perceptual abilities, hypersensitivity, focus on local detail or parts of objects at the expense of wholes, difficulties with generalization and abstraction, and lack of sensitivity to the context in which perceptions occur (Happé and Frith 2006; Mottron et al. 2006). Related to this autistic perceptual endophenotype is a characteristic insistence on sameness, as classically noted by Kanner (1943):

‘A situation, a performance, a sentence is not regarded as complete if it is not made up of exactly the same elements that were present at the time the child was first confronted with it. If the slightest ingredient is altered or removed, the total situation is no longer the same and therefore is not accepted as such’ (Kanner 1943: 246).

In other words, the child does not generalize across situations, disregarding perceptual differences that in typical development (TD) would not compromise
sameness relative to a particular description (e.g. something being a breakfast, a car, or greeting, etc., even though different foods are eaten, the car looks differently, or a different sentence is uttered as a greeting or to a different person). As Fay and Schuler (1980) insightfully captured this phenomenon:

‘If there is one pervasive theme in the study of the language of childhood autism it is the permanence of the initial learning situation. How can speech be brought into line with adult models if the only associations are first associations that are tenaciously stored and recycled as if they were cast in concrete?’ [pp.77–78, Italics as in original].

In Fay and Schuler’s terms, language comes to denote, but not connote, with words used more rigidly as proper names in context-specific ways. The resulting more ‘perceptual’ cognitive style can be paired with profound impairments in forms of cognition that are intrinsically social in a way that a particular sensory modality such as vision cannot as such be. This includes frequent problems with eye contact, gaze following, joint attention (JA) and shared intentionality. As noted, language is inherently a form of social cognition as well. Remarkably, however, language is rarely mentioned in this connection in the autism literature, where difficulties with communication and social interaction are standardly not regarded as linguistic difficulties but as causing these where they exist (e.g., Boucher 2012). Language, however, as argued here, mediates sociality of a particular type, and it is this very type that appears relevant to autism symptomatology. Language crosses between individual minds, while still conveying the potentially different contents of each of these. While the thought contents are individual ones, their truth is independent of the minds in which they occur. Language thus both mediates and transcends belief. It transforms cognition by providing a layer of information processing not dependent on perception, yet inherently integrated with it, enabling a new form of learning from communication as well.

Any lexical item of the substantive lexicon used in such communication, moreover, has general meaning. A parent calling something a toy refers to it as an instance of a kind. Drastic perceptual differences between toys become irrelevant when the issue is whether they are toys. With regards to that description, they will count as the same, no matter whether the toy of the initial learning situation was yellow or not. Language cannot, therefore, be processed without the abstraction from perceptual and contextual detail that creates such characteristic difficulties in ASD. This concerns events no less than objects and persons, since events, too, like a dog’s chasing a cat, can be made up of very different-looking and sounding dogs, cats, and chasings, while they nonetheless all fall under the linguistic description dog chasing cat (de Villiers 2014). At the same time, although words transcend perceptual context and content, any
linguistic utterance, despite involving general concepts, is always contextual too, in that no sentence can be fully understood without also grasping the speaker’s intention and knowing the situation in which the sentence is uttered. The sentence *Peter sent me a kiss* means nothing in isolation: Who are *Peter* and *me*? What does it mean to send a kiss? What kind of kiss? When was it sent? Why is this relevant?

If language has (i) reference as integrated with description, (ii) generality and also contextuality, (iii) it is inherently social and shared, and (iv) it is an ingredient in virtually any human social interaction, it is natural to expect that disturbances in the unfolding of this capacity can have profound consequences for a child’s apprehension of the world and cognitive development at large, in other words that language can be a causal factor. It makes sense in particular that a more perceptual, context-confined, rigid, and non-social cognitive style may result. Language should be a prime focus in theories of autism for this reason, quite apart from its practical significance as a prime reason for initial referrals and its role in predicting long-term outcomes in affected children (Rutter 1970).

### 3.1.2 Theories of ASD

Language had been such a focus until the 1980s, where language impairment was an inherent part of ASD diagnostic criteria. This ceased to be the case with the recognition that language levels can differ widely across the spectrum, with many children, teenagers and adults coming out as ‘language-normal’ according to standardized clinical measures and lacking a history of language delay. Since autism symptomatology without language impairment cannot be due to language impairment, non-linguistic cognitive deficits have to be invoked. But autism symptomatology lies by definition in the domains of communication and social interaction. Therefore, it logically follows that deficits in these domains in this case must concern communication and social interaction in non-linguistic senses of these terms. But while language might indeed dissociate from communication and social interaction – a global aphasic could communicate a lot using gestures and drawing, apart from socially interacting normally, while a psychopath or mass murderer could be deviant in communication and social interaction yet be linguistically normal – it does not appear that, in the case of ASD, the communication problems are mainly or only in the non-linguistic domains. On the contrary, people with ASD find various ways to communicate, though they are not always the ones we expect from linguistic beings; and idiosyncratic and socially inappropriate behaviours can have
communicative functions (Brady and Keen 2016). If problems of communicative and social interaction do involve problems in the social use of language, however, then such individuals can be ‘language-normal’ only if we define ‘language’ so as to exclude its communicative and social use from this definition, which is circular.

Be that as it may, none of the major current theoretical approaches give language a central cognitive role for understanding the neurocognition of ASD. Instead, these approaches can be grouped as revolving around hypothesizing deficits in one or more of three broad neurocognitive domains:

(A) Social cognition in a non-linguistic sense, particularly including mentalizing/mind-reading (‘theory of mind’, ToM: Baron-Cohen et al. 1985; Frith 2001) and joint attention (JA; Mundy 2016) deficits;

(B) Perceptual processing (‘weak central coherence’, WCC: Frith 2003; Happé and Frith 2006; enhanced discrimination paired with reduced generalization: Plaisted 2001; ‘enhanced perceptual processing’: Mottron et al. 2006);


According to the classical ToM hypothesis, people with ASD have difficulties in representing the mental states (beliefs, desires, etc.) of other people. WCC posits a processing bias resulting in failure to extract global form/meaning from perceptual inputs. EF is an umbrella term intended to denote a range of higher-order cognitive abilities necessary for flexible and adaptive action in the service of novel goals, including such abilities as planning, decision making, working memory, cognitive flexibility, set-shifting, impulse control and inhibition. As things stand, (A) to (C) have long histories by now and their respective virtues and problems have been widely discussed. What has not been discussed widely is their possible connections with language functioning. Such connections can be fruitfully explored for (B), since generality and abstraction are essential to the meaning of language and the information conveyed therein as noted: this is an elephant uttered by a parent in the zoo means that any animal looking like the one pointed at is an elephant. The question Do you want that dinosaur? could ask whether the child desires a particular kind of dinosaur. The command Don’t grab the food like that! asks to refrain from a specific action just undertaken, but also an infinity of other actions falling under the same general description. At the same time, reference in language can be to specific objects viewed as wholes, using descriptions identifying their natures and conditions of identity (e.g. car, not set of spinning wheels). For possible connections between language and components of executive functioning, see e.g. Akbar et al. (2013), Boucher and Mayes (2011), Paul et al. (2008). Here, however, I would like to focus on (A).
Language as present socially from birth and before not only establishes a social bond and selectively guides attention, but when developed, it is also one of our primary tools to figure out what people are thinking. Our conclusions on what they think cannot be separated from our understanding of what they say, and we use reasoning encoded verbally to generate hypotheses about what goes on in their minds. Does he think I lie? Might she want to come with me? Does she like me? Does he mean me? Language in this way takes us beyond what we could extract perceptually from facial expressions or behaviours, also serving as a corrective measure for perceptual inferences by counterbalancing them with verbal reasoning. Using language in these ways is not an optional extra, moreover, but language used in other ways (e.g. for behaviour regulation, as it primarily is in low-functioning ASD: Maljaars et al. 2011), is a clinical phenomenon.

(Explicit) mind-reading, moreover, is defined as a meta-representational ability, where one propositional representation (e.g. ‘I lie’) is embedded in another (e.g. ‘he thinks I lie’). This ability mirrors the process of subordination that is an inherent aspect of the grammatical process, in which phrases get embedded in others, first intra-clausally and then inter-clausally, giving rise to different grammatical meanings as a result. Sentences with clausal complementation of verbs ipso facto feature meta-representational content, since they encode thoughts about thoughts (or representations of these, as in reporting what someone said). With grammatical Person distinctions thrown in, the result is a meta-representational tool that not only covers one’s own first-personal mental states but also those of others, including what we think they might think from their own first-personal perspective. ToM, by contrast, does not as such entail either propositionality or a threefold Person distinction. Where the meta-representation that it necessarily involves would come from, if not from language, is also unclear. As a psychological construct, ToM remains described by its broad function and formally unspecified in terms of the generative mechanisms it is based on. How exactly does ‘representing the mental states of others’ work? Describing the lexical and grammatical structure and function of a sentence like He thinks I am lying would address this exact question.

In line with this perspective, language and ToM have long been shown to be developmentally linked, in both preschool TD children (de Villiers 2007), older children with ASD (Lind and Bowler 2009), and deaf children (de Villiers and de Villiers 2012). Training studies (in TD: Hale and Tager-Flusberg 2003; Lohmann and Tomasello 2003) as well as longitudinal studies (in TD: Astington and Jenkins 1999; de Villiers and Pyers 2002); and ASD (Tager-Flusberg and Joseph 2005) moreover provide some evidence that language may influence false-belief reasoning rather than vice-versa. A prime suspect for the linguistic component
that has the relevant impact on mentalizing has been clausal complementation (de Villiers and Pyers 2002; Paynter and Peterson 2010; Tager-Flusberg 2000; Tager-Flusberg and Joseph 2005; Lind and Bowler 2009). One could also point to a lexical difficulty in ASD with processing mental state verbs, i.e. that these ‘children do not understand linguistic terms referring to thinking, pretending, etc.’ (Baron-Cohen 1989: 586). Understanding such terms however appears to be closely linked to understanding how they grammatically function, namely by taking an embedded clausal complement describing the content of the thinking, pretending, etc.

A now widely accepted distinction, furthermore, is that between ‘explicit’ and ‘implicit’ mindreading. The former is slow-developing, cognitively demanding, and most importantly for our purposes here, verbal. It is the kind that ‘allows us to deliberate about mental states and to express our thoughts about mental states in words’ (Heyes and Frith 2014). Implicit mindreading, by contrast, of the type that is found already at 7 month olds, is a different and nonverbal mechanism, which operates rapidly, automatically, and efficiently, and unlike explicit mind-reading makes no demands on executive functioning abilities such as working memory and inhibitory control (Bull et al. 2008; Qureshi et al. 2010). On this two-systems account there is no single mentalizing system which is continuous across the developmental trajectory, until explicit ToM tasks are mastered. On Heyes’ (2014a, 2014b) specific view, the implicit system is no mentalizing system at all, depending instead on domain-general neurocognitive mechanisms that mediate automatic attentional orienting alone (see Butterfill and Apperly 2013, for another candidate). If so, there is only one true mentalizing system and it is verbal. In contrast to earlier nativist theories of mind-reading, moreover, which attributed mind reading to a genetically inherited, evolved and specialized mechanism, its acquisition may be due to a process of cultural learning in which language tuition plays a critical role (Heyes and Frith 2014). We ‘learn about the mind through conversation about the mind’, which teaches children ‘mental state concepts – what it is to “think” or to “feel” something, to be “happy” or “doubtful” – and gives them a format in which to represent these concepts’ (ibid.). Clausal embedding provides a crucial format in which such teaching and learning takes place.

In sum, while it is arguable that an underlying disturbance in language functioning is consistent with central aspects of all three main neurocognitive deficits posited in current theories of ASD, a promising and more specific hypothesis is that language mediates a species-specific form of reference based on general, lexicalized concepts used predicatively; and that without this mechanism properly developing at different levels of grammatical complexity, thought and experience will deviate in the autistic direction making mind-
reading more difficult in particular. This makes predictions for what kind of
linguistic anomalies we should find in the autism spectrum: they should con-
cern language as the mechanism as which it is described here, designed to
convert lexicalized concepts into referential acts straddling different levels at
which language serves to make sense of the world.

3.1.3 What studies of language in ASD to date reveal

It has often been noted that language difficulties in ASD do not merely involve
delay but also deviance, i.e. an atypical course of acquisition: language is
construed differently by children who face these difficulties (Eigsti et al. 2007;
Boucher 2012). According to recent estimates, a sizable 25–30% of children on
the spectrum never develop an idea of language at all (Norrelgen et al. 2015; Bal
(HFA), where there has been no language delay or where language has reached
normal levels by school age, language can remain affected when measured at
the level of discourse, pragmatics, and higher-order semantics. Thus children
with HFA without any language impairment as measured by standardized tests
can show such impairment in a story generation task (Norbury et al. 2014;
Banney et al. 2015). Even in people with a diagnosis of Asperger’s, language
can be abnormal behaviorally in prosody, volume, pragmatics, rate, and com-
prehension (Noterdaeme et al. 2010) and be processed differently in the brain
(Radulescu et al. 2013; Moseley et al. 2016). The same even applies to ‘optimal
outcome’ cases of adults no longer meeting diagnostic criteria for ASD (Suh et al.
the spectrum as emerging:

‘across group studies of the least affected to the most affected school-age children and
adults with ASD, with mainly quantitative differences in the degree to which any facet of
language is affected’.

The very existence of such a general profile makes it doubtful that this profile
would be caused by co-morbidities present only in a subgroup. In particular,
while the canonical linguistic profiles of ASD children with lower verbal abilities
and of SLI children can overlap in preschoolers at some level of description, they
come to diverge by school age (Geurts and Embrechts 2008). By this time, initial
phonological and syntactic difficulties shared between the groups appear to give
rise to the higher-order semantic and pragmatic difficulties more prominent in
the ASD group. More impairment in receptive than expressive language in the
same group is also atypical of SLI. A study of 120 children with ASD by Jarrold
et al. (1997) also found no evidence of different language profiles in any diagnostic subgroup on the autism spectrum. Eigsti et al. (2007), too, in a study of 16 children with ASD matched on lexical level and non-verbal mental age with children with developmental delays and typical development, found an even distribution of grammatical impairment across the ASD group. This impairment, unlike non-verbal IQ, also correlated with a cognitive measure, namely reference to non-present objects or events. There was no impairment in lexical competence, in line with other evidence that higher order receptive and expressive language is more impaired than single word vocabulary (Tager-Flusberg and Joseph 2003).

What then is the ‘ASD-typical’ language profile as stabilizing by school-age? Boucher (2012) identifies it as:

(i) Affecting receptive language more than expressive language, with productive syntax unimpaired, although the use of words and phrases can be idiosyncratic;
(ii) Articulation is largely preserved;
(iii) Morphemic anomalies present early can persist into adulthood;
(iv) Relatively normal performance on vocabulary tests, but higher-order semantic processing of words in context is anomalous.

Crucially, however, the terms used in this profiling are technical. It turns out that ‘syntax’ is here used to largely capture putting words in the right order, where the present approach predicts no difficulties. The term ‘morphemic anomalies’ is used to capture the use of ‘truncations, omissions, or substitutions especially of “closed class” words such as conjunctions, articles, or pronouns’ (Boucher 2012: 224). These are the hallmark of grammar on standard linguistic theory and the present approach centrally predicts difficulties here. They are critical to grammatical meaning involving deixis and Person in particular. As for ‘semantic processing’, this is affected primarily in higher-order domains, transpiring in the meaningful use of language at a grammatical level. Overall, this profile points in interesting directions with regards to the present proposal and it is certainly consistent with the idea of a disturbance in a language-mediated form of thought across ASD. In the remainder of this section I will try to make this point linguistically more specific, focusing on the notion of reference.

Among the earliest language disturbances identified by Kanner were echo-\v{e}lalic language (immediate or delayed), stereotyped or idiosyncratic language, language used as stimulans, neologisms, pronoun reversals, etc. (Kanner 1943, 1946). All of these are descriptively disorders in how language is normally used, namely referentially, using descriptions and Person distinctions appropriate for a hearer to identify the referent. Descriptively, the choice of a neologism, deviant
use of grammatical Person distinctions, use of language as stimulation or as a way of taking turns communicatively without saying anything, all disturb this process of reference as mediated by grammar in different ways. Related to this, Bartolucci et al. (1980) documented different but consistent patterns in the use of grammatical functors (inflectional morphemes) in ASD, which they interpreted as indicating a deficit in Person and time deixis, thereby highlighting deficits in the higher (definite/deictic/personal) ranges of the hierarchy of forms of reference described in Section II above. By now, experimental findings about disturbances in the grammatical Person system have also been replicated in several studies. Very recently, Shield et al. (2015) found that native signing children with ASD prefer to self-refer via their name-sign in a picture-identification task, where pictures depict either the experimenter or the child. In American Sign Language (ASL), personal pronouns are indexical pointings to self or other. They therefore clearly pick out their intended referents. If the cause of well-known difficulties with pronouns in spoken language in ASD was merely pragmatic, they should therefore disappear here. Moreover, there was no general pointing impairment, and the experimenter’s question (‘Who is this?’) contained an indexical point to the picture, thus modeling the very form that the answer should contain. Name-signs moreover are costlier to produce and less used than proper names in spoken languages in direct address. Despite this, the name-sign, a grammatically 3P form, was preferred, consistent with a pattern found earlier in hearing children with ASD in Jordan (1989), Lee and Hobson (1994) and Mizuno et al. (2011). The overall production of pointing signs was also strongly correlated with the level of ASL comprehension, as would be predicted from their intrinsically linguistic nature.

In a recent corpus study, Dascalu (2014) further investigated the different types and distribution of all occurring non-standard forms of self-reference in two French autistic children aged 4 and 5, respectively. These forms included (i) the use of ‘il’ (he) for ‘je’ (I), (ii) ‘tu’ (you) for ‘je’, (iii) own proper name, (iv) ‘il’ plus own name, (v) own name plus ‘il’, and (vi) null form (dropped subject). Of these, overwhelmingly, both children overused (i), i.e. ‘il’ (3P) for ‘je’. Crucially, the reverse – replacement of expected 2P or 3P forms of reference by the 1P pronoun ‘je’ – was not found (though it may be expected in cases of echolalia). This, together with the data from the naming studies above, documents an asymmetric Person shift away from 1P to non-1P forms of self-reference, with a preference for 3P- or non-personal forms. Crucially, on the other hand, self-reference as such is not affected: The child saying You want milk intending to self-refer is not confused over who wants milk. The difficulty lies with grammatical Person, not with reference or communication as such, in a non-grammatical or purely semantic sense. The significance of a reference deficit centered on 1P is
Further supported by (i) significant impairments in autobiographical narrative (which is 1P; Brezis 2015); (ii) better understanding of 3P rather than 1P false-beliefs (Williams and Happé 2009); (iii) use of 1P clitics as a distinguishing mark between an ASD and SLI cohort, who otherwise had very similar grammatical profiles (Durrleman and Delage 2016); and relative more difficulty with 1P perspective taking than 3P (Mizuno et al. 2011).

In their pronoun study, Lee and Hobson (1994) asked: Could the anomalous use of personal pronouns reflect not just linguistic competence but also the psychological experience of selfhood? The formulation is consistent with an un-Cartesian stance, which invites the thought that a disturbance in the grammatical Person system as reflected in personal pronouns is a disturbance of selfhood, precisely insofar as 1P forms of self-reference are essential to such human-specific forms of selfhood (Hinzen and Schroeder 2015).

While it is logically possible that disturbances affecting personal pronoun use is caused by a non-linguistic disturbance of selfhood, it transpires that pronoun misuse is the tip of an iceberg of problems with linguistic reference in general, however. Modyanova (2009) found, in a procedure where children with ASD had to act-out ‘now turtle touches a/another/the/that apple’ after another animal, Fishy, ‘had touched an apple’, that performance on ‘a’ and ‘another’ caused no problems for children with ASD, while acting out ‘the’ and ‘that’, which requires touching the same apple, did: the children were at chance. In no subgroup on the autism spectrum was performance on ‘that’ worse than on ‘the’, as a pragmatic account would have predicted. Difficulties with anaphoric definite noun phrases and pronouns have been highlighted in several studies of narrative in ASD (Baltaxe and D’Angiola 1996; Banney et al. 2015; Norbury, Gemmell and Paul 2013).

Disturbances of Person thus transpire in the context of a more general difference affecting 3P forms of reference as well, and primarily ones that are ‘high’ on the referential hierarchy. In line with this, long-standing data show that declarative pointing as such is impaired in ASD. Children with ASD use such pointing less than TD children or it takes anomalous forms (Goodhart and Baron-Cohen 1993; Camaioni et al. 2003). Hobson et al. (2010) found that the autistic children in their study could easily point, but only in this group were there participants who misused the deictic expressions ‘this’ and ‘here’ to refer to a location that was distal to themselves, or made atypical points with unusual precision, lining-up with an eye. Participants with autism were also less accurate in comprehending deictic terms, and ostensive referential signals. Hobson and Meyer (2005) in turn found that in a task where children had to communicate where on the body a sticker had to be placed, the autistic ones succeeded in the communicative task, yet never pointed to a place on their own body, as TD
children did. As noted, while pointing is often classed as a ‘non-verbal’ communicative behaviour (Tomasello 2008), it is highly correlated with language development and arguably the first appearance of a form of reference inherently linked to language development in humans. Communication for purposes of behavior regulation and social interaction, by contrast, is not so linked. Maljaars et al. (2011) in a study of 26 children with verbal low-functioning ASD, found that these children communicated overwhelmingly for behaviour regulation purposes rather than social interaction or JA; nonverbal children with ASD did so to an even higher degree. It does not seem implausible to suggest that contrary to what current diagnostic criteria suggest, children with ASD do not have a communication problem as such: their problem is in a human-specific form of communication, namely language. In line with this, Balboni et al. (2015) report that it is largely speech-related skills, not social interaction as such, which distinguished a group of low-functioning preschoolers from a matched group with neurodevelopmental disorders.

Several studies also found a predictive relationship between JA and development of language production in ASD (Mundy et al. 1990; Drew et al. 2007), suggesting that JA and language skills are developmentally linked. While declarative pointing and language use requires JA, the reverse is not the case: JA could be present together with imperative communication only, and attention as such, whether joint or not, does not entail any capacity for declarative reference. It would be difficult, therefore, to see our referential capacity as grounded in our JA capacity.20 Since newborns are exposed to language from birth, on the other hand, and have a preferential bias for speech over non-speech, and parental speech guides the baby’s attention even when visual skills are barely developed, it may be promising to regard language as providing a foundation on which JA skills in humans are built from the beginning, predicting deficits in JA when language does not fall into place normally.

Supporting this line of argument is evidence for basic disturbances in speech perception in ASD, which could have resounding effects for what contents are learned from communication in speech and the establishment of the deictic frame. Alterations in speech perception (Alcantara et al. 2004) and the preference of speech over non-speech (Klin 1991) have been documented and neurophysiological methods have revealed a ‘profoundly different stimulus-processing manner in autism’ partially specific to speech sound analysis (Kujala et al. 2013), which predicts subsequent effects on the processing of grammar and meaning, and hence learning and understanding. Speech (or sign), the vehicle of linguistic communication, which preferentially captures

20 I owe this observation to Joana Rossello.
newborn infants’ attention from birth (Vouloumanos and Waxman 2014), does not attract autistic children as much. Typically developing 6-month-old infants grasp the abstract role of speech in communication before having knowledge of words (Vouloumanos et al. 2014), which enhances their development, whether of language or thinking. By contrast, even siblings at high risk for ASD at 12 months do not pay the same attention to speech as typically developing children (Vouloumanos and Curtin 2014). Deficits of prosody are also common (Peppé et al. 2007).

Using positron emission tomography, Boddaert et al. (2004) reported under-activation of Wernicke’s area in children with ASD listening to speech during rest, as compared with a matched group of non-autistic mentally retarded children. In a striking novel research paradigm, Eyler et al. (2012) found deficient left hemisphere response to speech sounds and abnormally right-lateralized temporal cortex response to language in a large sample of sleeping 1-year olds later diagnosed with ASD, demonstrating that abnormal language processing is part of the ASD signature long before it can be formally diagnosed by standard means. In adults with HFA, Just et al. (2004) using functional magnetic resonance imaging reported atypical activation and connectivity patterns in language processing, with the left inferior frontal gyrus (Broca’s area) underactivated and left superior temporal areas (Wernicke’s) overactivated (see also Verly et al. 2014; Moseley et al. 2016). Stigler et al.’s (2011) meta-analysis interprets this pattern as indicating over-reliance on the meanings of individual words rather than the meanings formed during grammatical processing. The same authors note evidence for an unusual use of visual imagery in the processing of linguistic meaning (see also Mills et al. 2013). Mizuno et al. (2011) specifically reported anomalies in the processing of personal pronouns in adults with HFA, involving the functional collaboration of frontal (right anterior insula) and posterior (precuneus) nodes.

In sum, although extant empirical research on language patterns in ASD has for the most part not been driven by specific hypotheses on how language should matter to the neurocognition of autism, and language impairment is widely regarded as secondary to a primary impairment in social interaction, available findings are highly interpretable in terms of deviant language development failing to allow normal cognitive and social development.

### 3.1.4 Section summary

Language delays, unusual language use and absence of language in ASD have left little doubt that language is a key piece in the autism puzzle. But while a whole
range of reference anomalies in autistic speech have been found at the more grammatical end of the referential hierarchy discussed here, they have largely not been described in linguistic terms or regarded as primary. Against an independently motivated conception of language as mediating reference and yielding propositional forms of meaning embedded in a triangular deictic frame, the language anomalies found speak to a theory of ASD to which language would be basic. Of particular significance in this regard is an asymmetric Person shift from the grammatical 1P to non-1P, which may yield new and more objective insights into what has been described as a ‘selfhood’ disturbance. Moreover, a language capacity that does not properly develop will deprive us of the power of one of our prime mind-reading tools; but also of an adaptive combinatorial flexibility that makes action guided by thought rather than stimulation and frees us from restrictions to contexts and interests. With language being intrinsically shared and social, moreover, the lack of this tool will diminish an essential social bond that in typical development unfolds from birth and before.

In the final section of this paper I will make a case for a related conclusion in the case of the schizophrenia spectrum. It is worth recalling in this regard that SZ in fact was considered as the same disorder as autism for much of the twentieth century, with autism as its childhood-onset version (for current evidence and views on their interconnection see Ellis et al. 2016; Chisholm et al. 2015).

3.2 Language in schizophrenia (SZ)

3.2.1 The symptoms of SZ in their relation to language

Three so-called ‘positive’ symptoms are diagnostically essential for SZ according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (American Psychiatric Association 2013: 295.90):

(i) Hallucinations
(ii) Delusions
(iii) Formal thought disorder (FTD)

Of these, FTD is clinically a virtual synonym for ‘disorganized speech’ (speech that lacks coherence and is hard to follow), and language is thus written into its clinical presentation. That said, on a closer look, the other two symptoms crucially involve linguistic dimensions as well. Thus while it is well known that hallucinations neither need to be auditory nor verbal in SZ, auditory verbal hallucinations (AVHs) are clinically particularly prominent.
Already Bleuler (1911: 79–84) noted that elementary auditory hallucinations (e.g. hearing shooting or the wind blowing) were relatively rare. If they occur they were often interpreted as involving reference to the patient (e.g., the shooting occurs to rescue him), hence they really represent what are now termed referential delusions (‘the shooting occurs because of me’). Crucially, reference is nothing we can strictly speaking hear, and as we have argued here, it is a linguistic category inherently. Bleuler further notes, consistent with a recent large study (Baethge et al. 2005), that the most common auditory hallucinations are those of language: ‘music is rarely heard’. He also notes that many hallucinators ‘cannot tell whether they hear the voices or whether they only have to think them; they are “lively thoughts”, which are nonetheless still called voices by the patients themselves; then again they are “loud thoughts”, “toneless voices” (...)’ (Bleuler 1911: 90). What cuts across these phenomenological differences is precisely that language is perceived (whether or not sound is), in the sense that the thoughts in question have a linguistic articulation, come with a content that is given by this articulation, and that such thoughts/voices often appear as acts of linguistic communication to the recipient (even though the communicative situation is highly unusual). In short, AVHs are clinically central to SZ, and purely descriptively they involve a false perception of speech.

As for delusions, these occur formally as assertions, which a patient holds on to despite their often obvious falsehood, e.g. ‘The Mafia wants to kill me.’, ‘I am the Emperor of Antarctica’, ‘I have a power plant in my stomach’, or ‘I am Jesus’. These delusions couldn’t be the ones they are if they were not internally linguistically articulated in the way they are, containing the lexical concepts and grammatical relations that they do. Prior to delusions, delusional mood can occur, as when patients report that ‘Everything looks unreal’ or ‘Everything seems already set up for you like in a theatre’ (Fuchs 2005). These are distinguishable from delusions proper in terms of their linguistic form (neither of the two examples just given could clinically be classed as delusions proper, based on their form). Even for delusional moods a loosening of conceptualization (e.g. people are seen as if on a stage) is required rather than merely one of visual perception (‘theatre’ is not a purely visual concept). Disturbances in salience attribution (Kapur 2003) or lower-level perceptual deficits (Uhlhaas and Mishara 2007) occur and are linked to delusions, yet cannot mean that for a delusion to fully arise and to be identifiable as such, it does not have to take a linguistically articulated, specific propositional form – a perception as such could not be a delusion.

Based on this we can now further note that it is certainly not the words that are anomalous in delusional utterances (i.e. kill, the, Mafia, etc.). Rather, what is
anomalous is how the words are put together, i.e. their grammatical combinations and the meaning resulting from this (i.e. what events are referenced and which people are said to do what). Such utterances are linguistically normal only if we consider grammatical combination as having nothing to do with the meaning inherently resulting from such combinations. Consistent with this approach, sentences used in delusional utterances are not only unusual, but they do not actually mean anymore what they do when uttered in mental health. There the utterance ‘I am Jesus’ e.g. could easily be made, but it would have a metaphorical or ironic use, or if not, it might be the assertion of an actor who points to himself as acting on a screen. When uttered as a delusion, on the other hand, this sentence ceases to have any of these meanings. Rather, the patient is confused about which 3P form of reference (in this case, ‘Jesus’) identifies himself as picked out in 1P – a confusion in deictic anchoring: ‘and a unique failure of the constraints on the use of 1P mentioned in Section 2.8. In fact, we don’t know any more who the patient is referring to: himself or Jesus – a distortion of the deictic space. Similarly, in mental health, the assertion that ‘I have a power plant in my stomach’ could only be metaphorical; and if ‘The Mafia is trying to kill me’ is not a delusion, then the speaker is referencing an event in context conceivably describable as ‘trying to kill’ someone and as providing indirect evidence for such an assumption. In the delusional case, no such event can be identified as being appropriately referenced, and no police investigation is triggered. Moreover, while in mental health any such utterance can be negated, questioned, or embedded (e.g. ‘I sometimes think I am Jesus’), embedding necessarily deprives a proposition of its delusional status. Delusions are therefore necessarily unembedded utterances; and what cannot be embedded cannot be questioned, potentially explaining the status of delusions as unquestionable truths (Hinzen et al. 2016b).

In sum, while in delusions grammatical sentences are generated that have propositional meanings and normal uses in mental health, these come to lack such meanings and uses when they occur clinically. In this way, while language is as absent from diagnostic criteria in the case of SZ as in that of ASD today, it is virtually written into all the positive symptoms of SZ. It is relevant to negative symptomatology as well, as clinical terms such as alogia (poverty of speech) or poverty of content suggest. Hinzen and Rossello (2015) therefore propose a re-description of positive symptomatology as given in Figure 2. While FTD is primarily a disorder at the level of speech production, AVHs reflect anomalous speech perception, and delusions represent a disorder at the level of content.

A different question is how language as a neurocognitive variable could explain symptoms and hence contribute to a neurocognitive theory of SZ as well. To this we turn next.
3.2.2 Language and neurocognitive theories of SZ

People with SZ often exhibit a broad range of neurocognitive deficits, in such domains as attention, executive functioning, memory, language, social cognition, processing speed, and perception (Reichenberg 2010). Moreover, as Reichenberg notes, ‘any specific neuropsychological deficit in schizophrenia occurs in the context of a background of a very severe general intellectual impairment’, making specific deficits difficult to isolate. Strikingly, on the other hand, a significant proportion of patients with SZ have a normal neuropsychological test profile. Reichenberg et al. (2009) compared neuropsychological functioning across SZ and schizoaffective disorders and found the prevalence of normal test performance to range between 16% and 45% in SZ. All four diagnostic groups in this study (SZ, schizoaffective, bipolar, and major depressive) moreover had comparable neuropsychological profiles, with differences largely quantitative, making it difficult to associate such profiles with SZ-specific symptoms. There is also an explanatory problem: How or why should deficits in any of the above specific neurocognitive domains or else a general intellectual deficit give rise to ‘reality distortion’ symptoms such as hallucinations and delusions?

Figure 2: The three positive symptoms as a breakdown of the linguistic frame of thought (Hinzen and Rossello 2015). Human language organizes speech perception, production, and content into a co-dependent triangle. As the baseline of speech between the speaker and hearer breaks, symptoms result, which differ depending on which corner of the triangle is most affected.
Where evidence for a specific neuropsychological deficit has been found in
SZ, their empirical association with reality distortion symptoms has in fact been
tenuous. Thus ToM deficits have by now been widely attested in SZ and they
appear disproportionate to the general level of cognitive impairment seen in SZ
(Sprong et al. 2007). Moreover, it seems intuitively plausible in this case, as Frith
(1992) argued, that a difficulty in ascribing mental states to other people may
lead deluded patients to attribute faulty intentions to others’ actions. Some of
these might thus appear as acts in which an agent (say, the TV presenter)
communicates obliquely with the patient (a referential delusion); or agents
may appear to have malevolent intentions towards the patient, leading to
persecutory delusions. However, ToM deficits, while widely attested in SZ,
have not been found to either associate with reality distortion symptoms as a
whole or with delusions or hallucinations specifically.21 Doubts concerning the
plausibility of the explanatory connection arise as well. Why does the patient
not reject the mistaken beliefs in question? Why do they harden into delusions
shielded against pervasive counterevidence? Why does the patient seem unable
to appropriately evaluate what we tell him in response? How would the account
extend to delusions such as ‘I am Jesus’, which are not related to perceptual
events in the way that referential delusions are?

In a referential delusion, moreover, the problem is not in attributing mental
contents as such: not the content of what the TV presenter says as such is
misrepresented (e.g., a car accident took place), but rather how this particular
thought content is embedded in the deictic space that includes the patient, who
in this case happens not to be the TV presenter’s addressee (2P) or the 3P agent
in the reported event that he may delusionally think he is. But as noted in
Section 3.1.2, the ToM construct does not per se have anything to say about
grammatical Person distinctions – as such, it is a psychological construct
denoting the ability to represent mental states in others. The deictic dimension
that is critical to referential delusions therefore has to be added to the ToM
mechanism, in a way that was neither intended nor necessary when ToM was
first applied to ASD, where the issue was ‘mind-blindness’ (Baron-Cohen 1989)
rather than deictic confusion. Since language in its normal use also is a mechan-
ism for unlocking minds, allowing us to attribute contents as well as deictically

21 Ventura et al.’s (2013) meta-analysis (154 studies, combined \(N=7175\)) found that ‘the
correlations between reality distortion and the domains of social cognition ranged from
near zero to moderate’ (see also Freeman 2007; Walston et al. 2000; McKenna and Oh 2005:
Table 5.5, p.122).
anchoring them, it is again not clear which benefits the postulation of this as a language-independent mechanism adds to normal language functioning.\textsuperscript{22}

The obvious question, however, is why a theory hypothesizing a linguistic dysfunction affecting the thought process should fare better. The answer is, firstly, that as argued above, language dysfunction is virtually written \textit{into} the core symptoms defining their clinical presentation – so they are directly relevant to them, and an explanatory relation that makes one particular neurocognitive variable relevant to a particular symptom does not have to be worked out. As argued above, this relation to language is clear for FTD and AVHs and it follows for delusions the moment that we accept that it is intrinsic to any linguistic utterance to have a specific kind of propositional content, which delusions uncontroversially lack (they are not taken to be candidates of truth by hearers, and not reasoned with as such candidates by the patients). Delusional assertions thus fail to be a problem with language, only if we do not consider such content as an inherent aspect of language. Language without such content, however, is not language in the same sense, but a clinical condition.

The linguistic theory also does not face the specific problem of explaining how a neuropsychological deficit as seen in neurological disorders would give rise to a reality distortion. Rather, it is based on the hypothesis that language is a cognitive system mediating a species-specific form of rational thought. That form necessarily includes a notion of truth associated to sentences uttered in a specific deictic frame, which perception or associative learning does not afford. With a notion of truth goes a notion of reality (and of thought as independent from it), turning language into the cognitive system that precisely and apparently uniquely generates the sense of reality that we see distorted in reality distortion. A dysfunction of language is therefore a significant candidate for entailing a reality distortion, of which hallucinations and delusions are clinical manifestations.\textsuperscript{23}

\textsuperscript{22} Association with reality distortion symptoms is a problem that befalls another major theory of delusions, which appeals to probabilistic reasoning bias (‘jumping to conclusions’) (for discussion see Hinzen et al. 2016b). The ‘dysexecutive’ theory of reality distortion faces a similar problem (Dibben et al. 2008). While association of neuropsychological deficits with disorganization and negative symptoms are generally more promising, dysexecutive accounts of FTD are also far from conclusive (Dibben et al. 2008; McKenna and Oh 2005: Chapter 4).

\textsuperscript{23} While phenomenological accounts of SZ often emphasize perception of higher cognition as the primary locus of disturbance (Uhlhaas and Mishara 2007), some phenomenologists themselves, notably Merleau-Ponty (1964) in a work originally titled ‘The origin of truth’, have stressed that language is an independent cognitive principle correlating with our sense of truth, which cannot be subsumed by perception.
This account connects with the earliest conceptualization of SZ through Bleuler (1911), who identified a ‘disturbance of association’ with a characteristic ‘loosening of associations’ as being fundamental to the syndrome he defined. Association was a broad concept for Bleuler, who used the term ‘law of association’ as being essentially the same as ‘law of thought’, both of which he tested via word association experiments, i.e. linguistically (Jung 1919: 4–5). He did not, however, make the law of association formally explicit. On the un-Cartesian hypothesis, this mechanism is grammar. It is, I have suggested, the only known and formally characterized principle for systematically combining lexical concepts so as to result in propositional meanings. As noted in Section 2.4, this process integrates all cognitive abilities. A disturbance in this linguistic process could therefore implicate a disintegration of our cognitive capacities in their relation to each other, resulting in different symptoms depending on which aspect of the architecture of language in the brain is most affected.

If language, on the other hand, merely were an expressive device, then even if a thought disorder only manifested as language disorder, as FTD does clinically, this would imply nothing about its inherent relationship to language. In line with this, Titone et al. (2007: 93) express a widely accepted view when they note that ‘the thinking anomalies associated with psychotic conditions are not, fundamentally, speech or language disorders (...). Rather, when language is used in an idiosyncratic way, it represents the outcome of a deviant thought process’. But no independent evidence for this claim is here provided, and it clearly depends on a prior Cartesian conceptualization of language. An independent reason is required for endorsing the Cartesian conception, when language de facto is inherent to symptoms.

In short, at a theoretical level SZ may represent the *disintegration* of a faculty that, in the case of ASD, may never properly *develop* in the first place, as reflected in the different times of onset of the two disorders. As in the case of ASD, we will now review findings on language in SZ in the light of the account above.

### 3.2.3 Studies of language in SZ

Linguistic studies of language in FTD were invigorated through Chaika (1974), who suggested investigating speech production disturbances in FTD without the

---

24 See also Critchley (1964): ‘the causation of schizophrenic speech defect lies in an underlying thought-disorder rather than in a linguistic inaccessibility’ (p. 359); or Maher (1972): expressed language is ‘a mirror of the [disordered] thought’.
prejudice of regarding these as epiphenomena of a thought disturbance, i.e. as empirical phenomena of interest to the linguist in their own right. That, however, makes a language disturbance not necessarily explanatory for what clinically is a disturbance of rational thinking. Indeed, few clinical investigations of language in SZ have proceeded based on any particular predictions for how or why language functioning would matter to this disease. An independent linguistic theory is required that connects language and thought and makes predictions for the forms of disintegration of thought that we will see when language disintegrates. Un-Cartesian linguistics is such a theory. A core disturbance in how grammar organizes meaning would, on this view, in particular predict again that reference is distinctively disturbed.

Reference however has not so far been systematically investigated from a linguistic point of view in SZ. Instead, language function has been investigated at its various levels of internal organization, using standard distinctions from linguistic theory such as syntax, semantics, and pragmatics (e.g. Covington et al. 2005), with all of these regarded as separate, while reference necessarily straddles all three of these levels. Specifically, regarding syntax as understood in a purely formal sense, the finding has been a pattern of reduced complexity and errors (Morice and Ingram 1982; Morice and McNicol 1986). But such a pattern does not appear to be specific to FTD or other symptoms of SZ; nor would it explain the disintegration of the thought process seen in FTD. Focusing on FTD, Oh et al. (2002) found sentence-level (rather than lexical) ‘semantic anomalies’ to be specific to FTD, where the notion of such a semantic anomaly corresponds to the idea that, when contemplating a single sentence only, the hearer has no idea what particular utterances mean (e.g. The pond fell in the front doorway). Other authors have set their hopes on discourse-level deficits that might explain why speech in FTD can be so hard to follow. Rochester and Martin (1979) in particular capitalized on ‘cohesive devices’ that tie sentences together, particularly emphasizing pronouns, which in FTD patients showed significantly more instances of ‘unclear reference’. This finding was replicated (Harvey 1983; Chaika and Lambe 1989), and problems with pronouns specifically have been noted more widely to characterize SZ (e.g. Frith 1992: 99; McKenna and Oh 2005: 112–114), though formal studies are scarce. Barch and Berenbaum (1996), too, highlighted ‘incompetent references’ in FTD defined as impairment in demonstrative and personal reference, i.e. reference at its deictic end.

Going beyond pronouns, Docherty et al. (2003) found ‘confused references’ and ambiguous word meanings – a problem at the level of description – to be associated with FTD, while she found other referential problems across the spectrum and even in unaffected first degree relatives (Docherty et al. 2000, 1998). Docherty et al. (1996) established six types of ‘communication
disturbances’ that distinguished a SZ group from healthy controls. All of these, it transpires, suggest a problem with referential definiteness, either at the level of abnormal vagueness (indefiniteness, i.e. insufficient descriptive specificity when such specificity is required), confused references (several candidates for what is being referenced), missing information references (proper names or definite descriptions are used without prior mention of the referent), uncertain definitional meaning including neologisms and paraphasias, and unclear meaning due to structural unclarities or formal ungrammaticality. An anomalous lack of definiteness is almost a defining property of a subtype of FTD, ‘poverty of content’ (‘empty philosophizing’), which contrasts with the incompetent use of reference at the higher end of the referential scale, where reference is definite and comes with existence presuppositions. Note also that delusions are definite virtually by definition: ‘I might be the emperor of Antarctica’ or ‘Some religious organization is trying to kill someone’ are even formally not as such delusions, though they may express delusional moods; neither is ‘Some person has a power plant in their stomach’.

It thus seems fair to say that the notion of reference – and particularly problems with referential definiteness and deixis including pronouns (which tend to be definite) – has already played a significant role in thinking about language in SZ and FTD in particular, though this was often under other labels such as ‘cohesive devices’ or ‘communication disturbances’. The substrate of such disturbances however effectively is the grammar of noun phrases under their normal referential use, some of which simply have cohesion and communication as their natural functions. It also seems likely that the referential disturbances in question do not merely represent pragmatic problems in handling these referential devices, as such problems can typically be repaired in discourse. Thus I might say to you that ‘Peter is ill’, presuming that you know who I refer to as Peter, and provide further information when the presupposition is seen to fail. Irony, metaphors and metonymies can have similar effects, temporarily derailing the discourse and making it more diverting. When I say that Peter is a butcher, and I know he is in fact a surgeon, reference is not to a butcher but to Peter, who may in some ways be similar to a butcher. Problems in figuring out reference in these cases are not deep, unlike in autism, where they can be (Happé 1995). As the referential intention of the speaker gets clear, the problem disappears. A deluded patient’s referencing her seven husbands, on the other hand, will lead to a doctor’s puzzled question who these are, when she met them, how she thinks this is legally possible, etc., and answers to such questions will typically lead nowhere. The speech, it turns out, truly has no content in the normal referential sense and the referential intentions cannot be figured out or be made sense of.
The problem of referencing in delusions and FTD is thus likely deeper than a pragmatic one and points to a more fundamental problem in the referential use of language. This problem could further connect with problems of reference inherent to AVH, where the voice hearer becomes a grammatically second person or third person of the voice’s speech, when in fact he is no such addressee or topic, and he does not himself take part in the conversation with what is effectively a disembodied speaker. To illustrate the problem for FTD, Moya (1989) transcribed clinical interviews with 46 severely thought disordered patients, of which one opens as follows:

Patient: ‘They call me the Moorish queen. Me, I go with the Moors.’

Doctor: ‘Let’s see, you go with the Moors?’

Patient: ‘Yes, of course, the ones that are in the heavens. This is the pure truth.’

Later in the conversation, after a long pause, the patient reopens the conversation uttering out of the blue: ‘The moors are miners’. Later she clarifies that they are ‘Normans’; and ‘sell umbrellas’. The same patient does not know why she is in the hospital or who brought her there; she references herself in the wrong (male) gender under the name of the saint San Francisco Javier, which she claims was her name in a previous life, while she also claims to ‘be of the Highest’ (‘soy del Altísimo’), who upon being assassinated by someone transferred his soul to (or became – it is not clear) her. Meanwhile the doctor asks who ‘they’ are who she thinks brought her here; who the Moors are; where they are; how the Highest was assassinated and by whom and when; etc. Reading through such transcripts is to witness a continuous quest for what is actually being said or intended. We don’t know who or what is being referenced under the descriptions given, while at the same time patients typically have no insight that reference is not clear. The pattern depicts language use free-floating of referential and deictic anchors – a pattern distinctively different from the avoidance and rigidity of deictic distinctions seen in ASD. In his linguistic study of such severely thought disordered speech, Moya (1989) argued that problems in reference tracking typically arise with definites, proper names, and pronouns, which is where a requirement for reference resolution primarily arises, and in time and place deixis.

While misuse of pronouns and other definite nominals reflects failures of reference at the grammatical and high end of the spectrum, the disorder can affect reference at the descriptive or lexical end as well, as in the case of paraphasias and neologisms (e.g. _adjacentment_ in the place of _wall_, _drown_ in the place of _sink_ when referencing a boat, or _un bulle_ when there is no such Spanish word). In paraphasias, e.g. _John petitioned his friend for dinner_,
referencing can be intact (the referential intent can be clear), but description is loosening; in neologisms, it can be lost; in clanging, it can be chosen only because of a sound association with a previously mentioned word. These anomalies appear to arise primarily as words are used in discourse, i.e. referentially, and not in, say, naming tasks. In other cases, normal words are chosen, but they are combined grammatically with others in illegitimate ways, e.g. *The pond fell in the front doorway* (Oh et al. 2002). Tangential answers, too (like a patient replying ‘You didn’t bring your prosthetic legs’ when asked by a doctor how he is doing), can be captured by saying that the patient’s referencing is unconstrained by what the question references. Chaika (1974) captured the phenomenon of derailment by noting that any association can determine the topic of the next sentence uttered – hence discourse is not in fact organized according to topics, but lexically-associatively. ‘Topic’ is a grammatical notion, reflecting the notion of a ‘sentence subject’ as compared with the sentence ‘predicate’ (the ‘comment’). It corresponds to what can become of a referent as it outlasts the sentence in which reference to it is made. Referents that do not make it to topics thus could reflect a dysfunction of normal reference.

Another notion that has captured the attention of researchers is that patients with FTD exhibit a ‘failure to make use of context’. This would explain why they can show a bias towards the strong or preferred meanings of a word even when the grammatical context would normally make us infer that another meaning was intended (Chapman et al. 1964). However, Chapman et al.’s finding not only failed to be consistently replicated (McKenna and Oh 2005: 105–106), but it also was not related to FTD. Later, Kuperberg et al. (1998) provided indirect evidence for a problem with context. In this latter study, patients with FTD, more so than patients without and controls, were less sensitive to the context in which a word appeared, yet here the relevant notion of context was that of the grammatical context (‘co-text’), though the study was of comprehension, while FTD clinically primarily appears to be a production problem. This finding is equivalent to a disturbance of grammatical meaning – propositional meaning resulting from the processing of words in their grammatical frames, as opposed to lexically or associatively. In line with that, Kuperberg (2010) further analyzed deficits in the comprehension of patients with SZ under the notion of a shift in the ‘dynamic balance’ between lexical-associative processing of language and grammatical meaning, with a shift towards the former. On this model there are ‘two streams of processing, one drawing upon semantic relationships within semantic memory and the other involving the use of combinatorial mechanisms to build propositional meaning’. In mental health, semantic memory and grammatical meaning interact to build rational contexts in which new utterances and perceptions take place and are interpreted. The failure of grammatical control on meaning, accompanied by (or
maybe triggering) over-activity of stored semantic word-level associations, predicts disturbances in the use of context. As argued in Hinzen (2015), grammar defines a form of meaning that is non-contextual: the utterance *She’ll fly to Paris tomorrow* means what it does, irrespective of whether the speaker is smiling while speaking or wears red shoes – a matter of context which may well determine how we react to his utterance, but not its meaning as determined by its grammar. Where grammar fails to control meaning and hence to demarcate it from context, anything can become part of linguistic meaning and everything can become context, as the content/context boundary becomes blurred: Propositionality is lost.

Overall, half a century of investigations of language in SZ leave the impression that a principled notion is missing that would relate language to symptoms, but not that data are lacking that might support such a notion. In particular, in terms of problems in linguistic meaning, the problem is not lexical but grammatical – and it concerns the meaning that grammar is hypothesized to mediate on the un-Cartesian hypothesis, and more specifically the conversion of lexical concepts into referential expressions of various types, which may then disintegrate at different points in the process giving rise to different symptoms. A direct and testable prediction of this hypothesis is that reference will disintegrate more at the higher (more grammatical) end of the referential spectrum, with pronouns more affected than lexical nouns.

### 3.2.4 Section summary

Unlike in ASD, language anomalies in SZ are essentially written into the clinical definition of core symptoms in the case of AVH and FTD. In the case of delusions, the same conclusion depends on their propositional rather than perceptual nature and exploits the notion of a change in (or absence of) the (propositional) meaning that is inherent to language in its normal use. Moreover, any explanation of ‘reality distortion’ will need to invoke a cognitive principle giving rise to a notion of reality or truth in the first place. Language, in establishing a non-associative and non-probabilistic form of meaning in a triangular shape, is a prime candidate for this cognitive principle. This is by no means to exclude potential lower-level deficits in such domains as associative learning, perception, attention, or salience processing. However, in over half a century of investigations of language in SZ, rich evidence has emerged on specific disturbances of reference, largely in the higher regions of the referential hierarchy, but of a more fundamental kind than we saw above in the case of ASD. In particular, while self-reference as such was not affected in ASD, it clearly is in delusions; moreover, in FTD reference can disintegrate at a purely
third-personal level as well, losing its deictic anchors and a content-context distinction, while in ASD the disorder appears to push in the opposite direction, towards too rigidified forms of reference that lack the flexibility of normal language use in its ordinary deictic frame.

4 Conclusions

Human beings engage in acts of reference of a kind not seen in other species, using grammar and words as an inherent aspect of their peculiar mode of thought. Evidence reviewed here suggests a dysfunction of language at the level of its referential function in both ASD and SZ (for evidence for the same conclusion in Huntington’s disease, see Hinzen et al. 2017). Critical reflection has to be devoted to the prevailing assumption that clinical language dysfunction needs to be either language-specific or else be the secondary consequence of a primary ‘cognitive’ impairment, a disturbance of ‘thought’. A focus on the core process of converting lexical concepts into grammatical and propositional forms of meaning through grammar could contribute to defining and differentiating cognitive phenotypes that exist within our species.

Acknowledgements: I am grateful to Joana Rossello for discussion and comments, and especially Hans-Martin Gärtner for meticulous readings of this manuscript.

Funding: Research leading to this paper has been supported by the grants ‘Language and Mental Health’, AH/L004070/1, and ‘Un-Cartesian linguistics’, AH/H50009X/1 awarded by the Arts and Humanities Research Council, UK, and the grant ‘Language, Deixis, and the Disordered Mind’ (FFI2013-40526-P) awarded by the Ministerio de economia y competitividad, Madrid.

References


Perspectives from developmental cognitive neuroscience (2nd ed.). Oxford: Oxford University Press.


Box 1: The intensionality argument

1. Thought involves *descriptions*, of which the speaker may not know that they apply to the relevant referent.
2. Thoughts become different thoughts with any *change* in these descriptions, keeping the referent constant.
3. Any mechanism generative of human thought therefore needs to generate the exact right descriptions.
4. Language is the only such mechanism known.

Box 2: One *significatum* and two modes of signifying: Illustrating grammatical meaning

A. “Mary smiles”
B. “Mary’s smile”
Caption: A significatum (what a representation is about) does not determine how we refer to it, via a Noun Phrase (NP), as in A, or via a verb phrase (VP) embedded in a sentence, as in B. This grammatical difference is not accompanied by a lexical one but entails a meaning difference: In (A) an event is being referred to with Aspect and Tense and embedded in a proposition with a truth value, while in (B) reference is to an object, with no truth value entailed.