

Beyond triadic communication

A three-dimensional conceptual space for modelling intersubjectivity

Max van Duijn and Arie Verhagen

Leiden University, The Netherlands

Coordinating different viewpoints is an essential part of human interaction. Languages have evolved conventional ways of supporting this process: many linguistic items are somehow involved in viewpoint management, ranging from morphological elements and lexical units to grammatical constructions and narrative patterns. In this paper we propose a conceptual model for analysing how particular instances (or combinations) of such linguistic items can be used to coordinate the viewpoints of signallers, addressees, and third parties involved in an interaction event. In essence, our model augments Langacker's (1987) "viewing arrangement" through the addition of a third dimension to the existing two. We discuss the details of our model using a range of examples from spoken discourse, newspaper articles, and literary fiction, and end by placing it in broader discussions on human social cognition.

Keywords: viewpoint, perspective, intersubjectivity, polyphony, viewpoint packages, framing adjectives, social cognition

1. Viewpoint management and intersubjectivity

Human communication is characterised by an ongoing development of perspectives and perspectives-on-perspectives, a "polyphony" of voices making themselves heard and interacting with each other (Vandelanotte 2007, among others; Bakhtin 1981). Not only do we share and coordinate our own inner life with that of the people we interact with, but we also constantly make implicit and explicit reference to the perspectives of others who may or may not be present at the time of speaking, or who may even exist only in the imagined worlds of thought and fiction. This suggests that many words and grammatical constructions (as well as other meaningful elements, such as gestures and facial expressions) in a stretch of

discourse must be somehow involved in viewpoint management. Or, paraphrasing Verhagen (2005:4): the assessment of and coordination with others' mindstates is such an essential part of human interaction that it is to be expected that languages have, over the course of their history, developed special conventionalised signals to support this function, in line with Du Bois' (1985:363, 2003) claim that "grammars code best what speakers do most".

As such, the idea that most language usage entails viewpoint coordination is recognised by many linguists and narratologists; however, focus has mostly been *either* on how signallers and addressees mutually coordinate their perspectives (in the linguistic tradition; see e.g. Langacker 1991, Sweetser 1990, Verhagen 2005, Du Bois 2007), *or* on how third-party perspectives are represented (in the narratological tradition; see e.g. Fludernik 1993, Vandelanotte 2009, Hühn, Schmid & Schönert 2009).¹ In this paper we discuss a conceptual model that integrates these functions all at once, instead of approaching them as distinct phenomena, thereby capturing the polyphonic nature of discourse – giving it a schematic face, as it were. Our starting point is the earlier version of such a model used in Verhagen (2005), which in turn was based on Langacker's (1987:139) "viewing arrangement". In short, we propose adding a third dimension to the existing two. The three dimensions in our updated model, referred to as (x), (y), and (z), correspond to three different types of cognitive coordination and linguistic signalling. The (y)-axis is used to highlight the relationship between interlocutors and objects that they jointly attend to, for example: *that football player*. The (x)-axis indicates negotiation of epistemic stances between interlocutors, such as *that football player is great*. The (z)-axis deals with the negotiation of the degree to which objects of joint attention are considered from the perspectives of third-party discourse participants, as in *she thinks that football player is great*. Here, *she thinks* is a signal used by the speaker to invite the addressee to also coordinate cognitively with a third party. Throughout this paper, the merits of adding a third dimension to the existing model of the construal configuration will be discussed in the light of various examples and existing approaches to intersubjectivity, viewpoint management, and epistemic stance marking.

1. An exception is Dancygier's work (e.g. 2012): her approach also integrates views from linguists' interaction models with narratological insights regarding speech and thought representation. See also several contributions to Dancygier, Lu, and Verhagen (2016). A theoretical model combining both is that of the *Basic Communicative Spaces Network* (BCSN) model proposed by Sanders et al. (2009); see footnote 10 for a suggestion how this model may be linked to the one proposed in this paper. Clark (1996) incorporates other participants in his model of communication, but these are all on the "addressee"-side, viz. different kinds of hearers.

We also introduce a graphic representation of our model in this paper. This representation *can* be used for schematically representing individual viewpoint configurations as prompted by particular linguistic items, and below we consider it illuminating to do this a few times when explaining the details. However, it is important to note that the purpose of our updated model is not primarily to introduce another practice of drawing schemas, but rather to make a general point about the structure and working of intersubjective interaction using language.

In the concluding section, the points made using our model will be discussed in the light of broader discussions on the relationship between language and *mindreading* (or *theory of mind*), the ability to form understandings of the inner lives of others in order to make sense of the social world around us. Our suggestion is that a better conceptualisation of the structure of intersubjective communication should eventually contribute to our wider understanding of the working, development, and evolution of intersubjective thought and human social cognition.

2. A three-dimensional conceptual space

2.1 Dyadic and triadic communication

Communication in non-human animals typically involves a sender producing some observable behaviour (the “signal”) that increases the likelihood of a receiver responding, i.e. behaving, in some particular way – for example, a bird signalling to a competitor to stay away from its territory. When the benefits of such a pattern of linked behaviours outweigh the costs for both senders and receivers, a (relatively) stable communication system may emerge. Thus, most non-human communication is about “regulating and assessing the behavior of others” (Owings & Morton 1998: i). At this very basic level, the conceptual space needed to characterise communication is one-dimensional: no other dimension than that of the sender-receiver relationship is necessarily relevant to characterise a signal and its causes and effects. In the words of Tomasello (2008: 23), animal communication is mostly “dyadic”: by far the majority of cases can be explained in terms of regulating others’ behaviours without having to take into account attention (let alone *joint attention*) to any objects of reference.

In contrast, human communication is prototypically “triadic” (Tomasello 2008: 23), as it is by default *about* referents in the (shared) world outside of the communicators and their interaction. Following this idea, the conceptual framework needed to characterise normal human communication should thus be at least two-dimensional: apart from the relation between the communicators, the

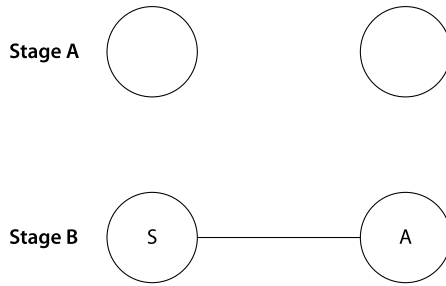


Figure 1. Schematic depiction of “dyadic” communication as found in non-human animals. In Stage A there is no communication between the two depicted subjects (circles). Stage B shows the situation in which the sender (S) transmits a signal to the addressee (A) in order to influence the latter’s behaviour, thus establishing a one-dimensional communicative relation. (Possibly, this induces a response signal: The subject on the right then becomes S and the subject on the left A.)

relation to the world must be taken into account to characterise signals and their causal connections. In other words, at the heart of interaction using language lies joint attention to some object of conceptualisation (person, event, relationship, etc.) and negotiating a particular stance towards this object.²

In simple terms, the two-dimensional conceptual space depicted in Figure 2 makes a distinction between the *intentional* aspect of language, its capacity to be *about* some object in the world, and the (*inter-*)*subjective* aspect of language, according to which sender and addressee negotiate a particular stance *towards* this object. These aspects are depicted by the (*y*)- and (*x*)-axes, respectively. It is basically equivalent to Du Bois’ “stance triangle” (2007:163), the corners of which are defined by a Subject₁ and another Subject₂, aligning with each other in evaluating some Object and positioning themselves with respect to it – clearly also a two-dimensional conception of human communicative interaction.

2. There are some instances of communication in non-human animals where functional reference to objects in the ‘outside world’ does seem to play a role; a well-known case is that of vervet monkey calls referring to different types of predators (see Seyfarth, Cheney & Marler 1980, Dennett 1987: chapter 6). Conversely, humans also regularly engage in purely dyadic communication, such as greetings (*Hi!*). However, as Owings and Morton (1998:211) argue, functional reference in animal communication should not be analysed as providing information about entities in the world, since it would confuse short-term with long-term causation; objects such as a snake in a snake alarm call should be seen as “long-term validators of the signal’s utility”; not as real-world correlates of signals which are causally involved in the receiver’s response to the signal. In human communication, however, triadic communication *does* prototypically involve real-world objects of joint attention, both as causes and as (intended) effects.

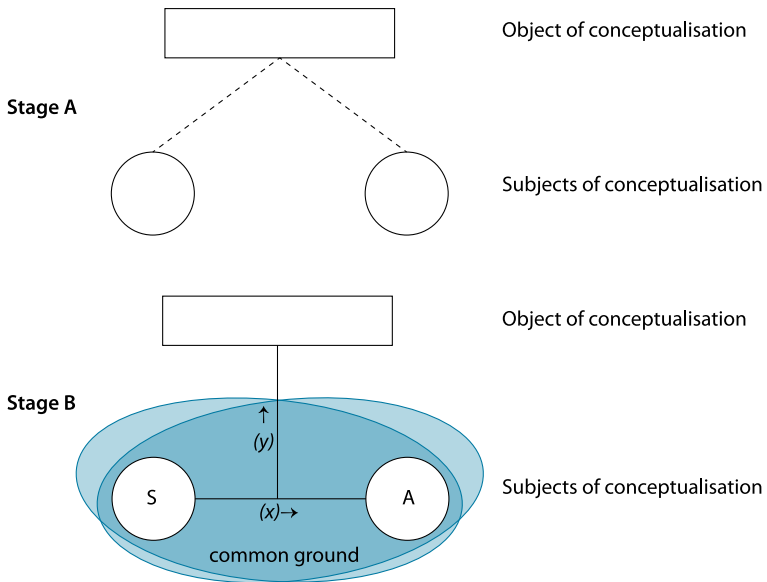


Figure 2. In Stage A, where no communication takes place, the two subjects (circles) both have their own views (dashed lines) on some object (rectangle). Stage B depicts triadic communication: The signaller/speaker (S) and addressee (A) both assume a set of shared beliefs (the overlapping part of which is the *common ground*; cf. Clark 1996) and subsequently negotiate how the common ground should be updated with respect to the object. As will be discussed below, the used signals typically reflect both aspects of and operations on the relationship between S and A (the (x)-axis) and on the relationship between the common ground and the object (the (y)-axis)

The field of cognitive semantics also embraces the idea that the proper characterisation of language use requires a two-dimensional conceptual framework, but its background, history and contents differ somewhat from the biological and psychological considerations presented so far. The cognitive view was explicitly designed by Langacker (1987) in opposition to so-called “objectivist” approaches to semantics, which held that meaning in natural language could be fully characterised in terms of no more than its relation to the/some world (its contribution to “truth conditions”). Objectivist semanticists were thus precisely ignoring the *perspectival*, subjective dimension, i.e. the (y)-axis in Figure 2. This axis is indispensable in Langacker’s view, since he claims that different “perspectival construals” are just as inherent components of linguistic meaning as reference is. Objectivist approaches to language thus in a sense also assume a one-dimensional conceptual framework for the analysis of meaning in natural language: they see all of meaning as pertaining to entities, their properties, and the relations they participate in *within* the object of conceptualisation (the/some world), without a role for

the orthogonal subjective dimension that is indispensable in Langacker's cognitive approach.³ Verhagen (2005) extended Langacker's model by including a systematic distinction between the viewpoints of the sender and addressee in order to bring out the fact that construal is not (just) a matter of a single viewpoint (*subjectivity*) with respect to some object, but one of mental coordination between signaller and addressee with respect to an object of joint attention (*intersubjectivity*). In hindsight, we can say that the framework proposed by Verhagen (2005:7) represents a merger of the biological and cognitive-semantic views of human communication.

What we will argue now is that a proper characterisation of viewpoint management in discourse (and of linguistic elements supporting viewpoint management) requires recognition of a third dimension. We will begin with a single case that presents a problem for two-dimensional models, and argue how the addition of a separate dimension, relating the present communicative situation to other ones, provides a straightforward solution. Subsequently, we will show that our updated model also provides a very natural framework for the analysis of other items and viewpoint configurations.

2.2 Speaker commitment and viewpoint embedding: Dalabon and English

Consider the following utterance in the Australian language Dalabon (Example (1)) and its English translation (Example (2)), as suggested by Nicholas Evans (class lectures 2009, brackets in original):⁴

- (1) *Ka-h-kangurdinjirrmî-nj yangdjehneng*
 3SG-ASS-GET.ANGRY-PSTPERF SUSPENDED COMMITMENT
bûrra-h-marnû-dulu-djirdm-ey
 3DUHARM.SUBJ>3SGOBJ-ASS-BEN-SONG-STEAL-PSTPERF

- (2) 'He got upset [because] [he thought that] the two of them had stolen his song.'

3. Of course, many semantic distinctions do pertain to all kinds of phenomena in the object of conceptualisation. There are all kinds of mappings between structure in the world and structure in languages, but as these are not directly relevant to our concerns in this paper, we will continue to represent the object of conceptualisation as a simple rectangle, without paying attention to any internal structure.

4. We thank Nicholas Evans for permission to use this example in this context. See Evans (2010) for more examples of elements for viewpoint coordination (esp. ch. 4), and for the glossing method used.

Because and *he thought that* are inserted in the paraphrase by Evans. We will briefly discuss the causal marker *because* in note 6, but focus on the insertion of *he thought that* in detail first. The lexical unit *yangdjehneng*, glossed by Evans as “Suspended-Commitment”, does not literally mean ‘he thought that’, but rather conveys the message: ‘I, speaker, am not committed (to what I am going to say now)’. A paraphrase closer to the original expression is thus (3):

- (3) ‘He got upset [because] [I, speaker, am not committed to this:] the two of them had stolen his song.’

At first sight, it may seem remarkable that Evans renders the lexical unit that functions as a marker of suspended commitment with a complementation construction in English – are the two indeed equivalents? To illustrate that, in an important sense, they are, consider the differences between the more idiomatic translation in (2) and the more literal one in (3). The absence of *he thought that* in (3) does not mean that *he* no longer had the thought that *the two of them had stolen his song*. In fact, awareness of the information in the second clause is equally implied in (2) and (3); if *he* had not had that thought, the stealing of the song could not have caused him to be upset. In both the Dalabon and English versions, the speaker invites the addressee to view the information about *the two of them* having stolen the song in relation to the mental state of a third party, namely, the *he* at the beginning of the sentence. However, there are differences in the degree to which this is accentuated and elaborated: the coordination of a third-party perspective⁵ is significantly more pronounced in the idiomatic English translation in (2) compared to the Dalabon original (1) and its paraphrase in (3).

With respect to the speaker’s perspective, the Dalabon element *yangdjehneng*, the English phrase *I, speaker, am not committed to this*, and the idiomatic pattern of sentence complementation using the stance verb *to think* play a similar role as well. All three negotiate an epistemic stance of the speaker towards parts of the presented content: in (1), (2), and (3), the speaker does not assert as true that *they had stolen the song*. And here, too, the difference is in the accentuation and elaboration: in (1) and (3) the tempering of commitment by the speaker is realised ‘on stage’, whereas in (2) this remains implicit.

Thus, in both the Dalabon and English versions, the speaker invites the addressee to view the information about the two of them having stolen the song

5. In this paper we use the terms *perspective*, *viewpoint*, and *point of view* interchangeably when referring to the general vantage point of a person’s thoughts and perceptions. We will use the terms *mindstate* and *intentional state* interchangeably to indicate *specific* thoughts a person can have. For instance, a scene can be ‘rendered from someone’s perspective’, whereas if ‘someone thinks that the house looks scary’, this is described as a mindstate. Assessing someone else’s perspective or mindstate will be referred to as *mindreading*.

from the perspective of a third party, and reduces her own commitment to this information. The differences concern the degree to which this is accentuated or profiled. In English, the embedding of the information in a third-party perspective is accentuated by means of a particular syntactic construction (complementation) and a particular matrix predicate (*thought*), while the speaker's reduced commitment remains more implicit. The Dalabon version, in contrast, profiles precisely this reduced commitment with the element *yangdjehneng*, while the third party's relatively higher degree of responsibility remains more implicit. In short, these conventional ways, in these two languages, of distributing responsibility for a piece of information over the speaker and another party are each other's mirror image: what is explicitly 'put on stage' and what 'is left to inference' is so to speak reversed. But the totality of what is communicated with these structurally very different expressions is very much the same, in particular the connections between different relevant viewpoints.⁶

Can both the similarities and the differences between these expressions be stated in a single analytic framework? Clearly, this is desirable, but when we try to do so by using the two-dimensional model of triadic communication in Figure 2, it soon becomes clear that this requirement cannot easily be satisfied. If the function of the element *yangdjehneng* ('I am not committed to this') is straightforwardly characterised as the speaker signalling to the addressee ((*x*)-axis) what his stance is towards ((*y*)-axis) the object of conceptualisation (i.e. 'the two of them had stolen his song'), the associated heightened responsibility for this view of the third party (*he*, the one who got upset), is necessarily left out. The reason is that this third party is only present in this model as an element of the situation being talked about, as an *object* of conceptualisation, and not as another *subject* taking a view on this situation.

Conversely, the representation of the complementation construction in (2) (*he thought that...*) does not work very well conceptually in a two-dimensional frame-

6. The phenomenon of marking explicitly only some aspects of what is to be conveyed is, of course, not limited to viewpoint expressions; on the contrary, it is quite general and well documented for various conceptual relationships, including causality (see, e.g., Verhagen 2005). It should be noted that an analysis similar to the one given above applies to the pair (i) *He got upset; the two of them had stolen his song* and (ii) *He got upset because the two of them had stolen his song*. The conceptual representation of both (i) and (ii) contains a causal relationship (otherwise no coherent interpretation seems possible), but this is only marked explicitly, 'on stage', in (ii). The difference between the Dalabon and English idiomatic ways of expressing both viewpoints and causal relations can be characterised as a difference in the available tools, and in the conventional rules for using them in the different languages. See also Wilkins' (1986) discussion of "particle/clitics" for criticism and complaints in Aranda, another Australian language, and his argument that these encapsulate "culture specific modes of thinking" that become clear when their use is explicated.

work either. In (2), *he* is not an object of conceptualisation: we are presented with what he thought, so he is at least a kind of subject.⁷ But at the same time, *he* should not be seen as a subject of conceptualisation *in the same sense as the speaker and addressee* either. After all, the negotiation of a stance towards the object of conceptualisation takes place between speaker and addressee, meaning that the speaker can, as a part of this negotiation process, *invite* the addressee to take the perspective of a third party on some aspect of the object of conceptualisation into account.⁸ However, perspective can never shift completely to this third party in the course of the interaction event being modelled (cf. the way it can jump from one character to another in a novel). In other words, the view of the third party *he* can be instrumental in the speaker's and addressee's negotiation of a stance towards the object of conceptualisation, but *he* is himself not a participant in this negotiation process, i.e. in the communicative situation. All in all, the common problem when representing the sentences (1), (2), and (3) seems to be that in a two-dimensional conceptual model of communicative interaction, third-person conceptualisers can only be situated either at the level of the object of conceptualisation or at that of the speaker and addressee,⁹ while in fact they normally belong to neither.

We therefore propose to treat other subjects of conceptualisation not as additional entities in the two-dimensional space, but as implying the addition of a third dimension, which links third parties exhibiting intentionality towards the relevant object of conceptualisation to the dimension of the negotiation process between speaker and addressee. The basic idea is captured in Figure 3:

7. The subject of *he got upset* does refer to an element of the object of conceptualisation. The fact that the second *he* is part of a complementation construction turns it into a reference to a viewpoint coordination device (a third-party conceptualiser). We do not at this point commit to a general theory on the conditions under which reference to mindstate functions as a viewpoint coordination device or an object of conceptualisation. However, our interest here is to make the claim that these two options are irreducibly distinct.

8. An anonymous reviewer of this paper raises the intriguing and important question as to when reference to a mindstate functions primarily as a tool for viewpoint coordination, and when as primarily part of an object of conceptualisation. If such a mindstate constitutes (practically) the entire conceptual content of an independent clause (e.g. *Then she knew*), then it functions as an object, and enters into a coherence relation with other event descriptions in the discourse. When it is the matrix clause of a complementation construction, the vast majority of cases function as a viewpoint device, but depending on the context, construal as an object is occasionally possible (Verhagen 2005: 112–113). Such a construal can also be supported by certain linguistic cues, such as stress and adverbial modification (Langacker 2014: 61–65). Thus it is certainly conceivable that there is no straightforward, definite answer to the question, but the issue definitely deserves further investigation. See also Verhagen (2008a, 2008b).

9. In Verhagen (2005: 106) this side of the dilemma was chosen to represent the relationship between first- and third-person conceptualisers. We can now see that this was an artefact enforced by the limitations of the two-dimensional model of intersubjectivity employed there.

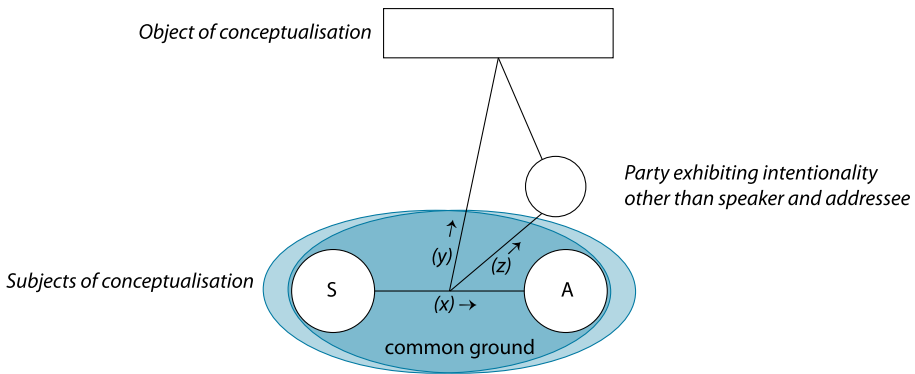


Figure 3. The three-dimensional conceptual model of interaction featuring a non-speaker, non-addressee subject of conceptualisation

We conceive of the third person represented in Figure 3 as a subject of conceptualisation in exactly the same way (i.e. with the same cognitive capabilities, including intentional reasoning) as the speaker and the addressee. Moreover, the object of conceptualisation for this subject is (at least in part) the same as the one that the speaker is inviting the addressee to consider, capturing the idea that the speaker presents the situation to the addressee *from a third-party perspective*.¹⁰

This basic model provides the conceptual space to mark precisely the similarities and differences between the Dalabon and English viewpoint items discussed above, regardless of the fact that they belong to completely different language systems. A graphic representation of the relevant components of these expressions can be found in Figures 4 and 5, respectively.

10. The logic of the problem for which we introduce this third dimension here is the same as that of Sanders et al. (2009) when they consider the question of how to represent the interpretation of the subjective Dutch causal connective *dus* in a third-person narrative context (*Jan zag dat hun lichten uit waren, dus ze waren niet thuis*, 'John saw their lights were out, so they were not at home', presenting their not being at home as a conclusion to the addressee, via John's point of view). Their solution, using their *BCSN model* (see footnote 1) in a way also resembles ours: they introduce a series of connected mental spaces into the content domain which are all absent in the basic communicative situation (compare their Figures 3 on p.33 and 7 on p.41). To us, this is an indication that the *BCSN* approach is compatible with ours in terms of the insights being captured. However, we consider it an advantage of our model that it explicitly distinguishes the dimension of different subjects of conceptualisation coordinating with respect to the same object of conceptualisation: in the *BCSN* model, the latter has to be represented in each relevant mental space, with linking lines between them, which complicates the representation (in two dimensions) considerably, compared to the 3D-model. Moreover, our model allows for a straightforward link, showing differences and similarities, to the analysis of communication (including animal communication) in general.

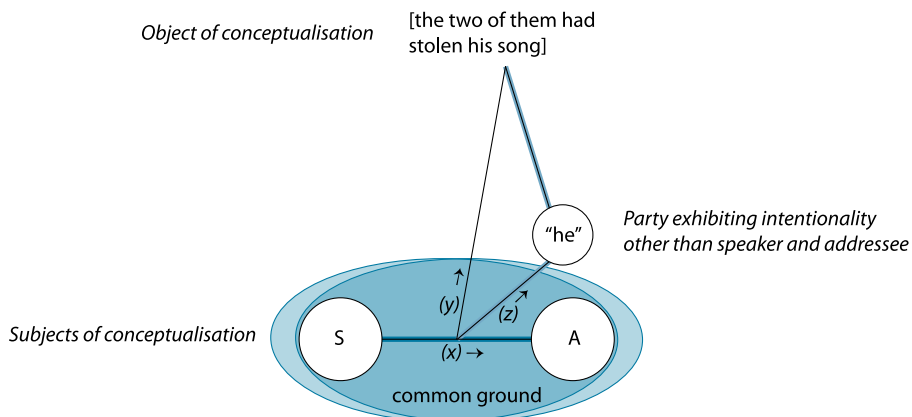


Figure 4. Schematic depiction of *yangdjehng* from sentence (1) using our 3D conceptual model. This Dalabon element operates explicitly on the (x)-axis, indicated with a dark blue line, and implicitly on the (z)-axis, indicated with a light blue line

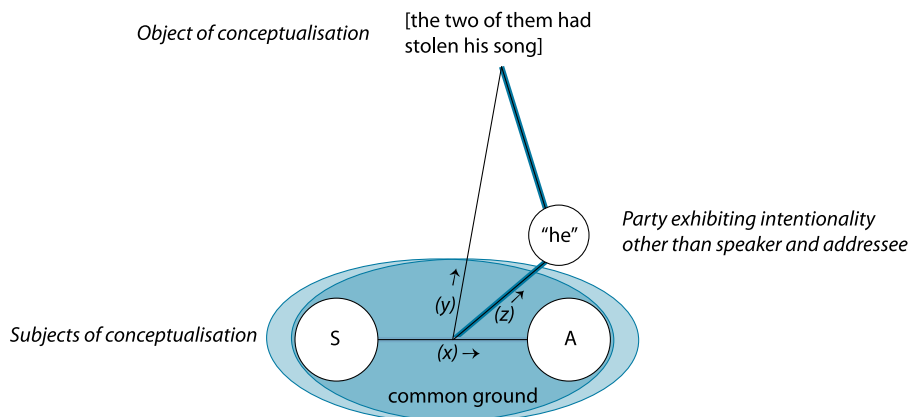


Figure 5. Schematic depiction of *to think + complement* from sentence (2) using our 3D conceptual model. This English element operates implicitly on the (x)-axis, indicated with a light blue line, and explicitly on the (z)-axis, indicated with a dark blue line

As in the two-dimensional model, the negotiation of epistemic stance performed by both the Dalabon and English elements is highlighted along the (x)-axis between S(peaker) and A(dressee). In Figure 4, this axis is marked with a dark line, indicating foregrounding of the speaker’s epistemic stance by the Dalabon element *yangdjehng*, glossed as ‘I, speaker, am not committed’. In Figure 5, it is the (z)-axis that is marked with a dark blue line, indicating that the English complementation construction with *to think* foregrounds the third party’s mindstate, while the (x)-axis is marked with a light blue line, indicating that this does impact

upon the negotiation of epistemic stance between S and A – in a less pronounced way than the Dalabon element does.

What is new in Figures 3, 4, and 5 compared to the two-dimensional version in Figure 2 is the (*z*)-axis connecting the (*x*)-axis to a third party, in this case the person referred to using *he* and *his* in Evans' translation in (2). On this axis, the reverse pattern obtains with regard to profiling: whereas in Dalabon this third-person perspective is only implied, indicated by a light blue line along the (*z*)-axis in Figure 4, in English it is explicitly realised 'on stage', indicated by a dark blue line on the (*z*)-axis in Figure 5. Thus, thanks to the additional (*z*)-axis, we now have a single format for representing that both the Dalabon and the English versions of the utterance invite the addressee to take the third person's perspective on the matter being talked about into account (i.e. the – actual or imagined – stealing of the song by *the two of them*), and that they do so in different ways, by highlighting what parts of the configuration are linguistically marked in each language, and which are implicit, but inferable.

2.3 The general model

When we first introduced the three-dimensional model, we stated (below Figure 3) that the additional intentional party is a subject of conceptualisation whose perspective is instrumental in the speaker's and addressee's negotiation of how the common ground should be updated with respect to an object of joint attention, without himself being a participant in this negotiation process. However, such third parties may themselves be represented by the speaker as being involved in another communicative interaction event, and, in fact, such parties may be talking or thinking about yet another interaction event. Thus, we may in principle expect to encounter more elaborate constellations of several subjects all in some way considering the same object of conceptualisation from different viewpoints, and affecting (more and less mediated through the viewpoints of others) the negotiation between S and A of epistemic stance, attitude, etc. Such a constellation is depicted in Figure 6.

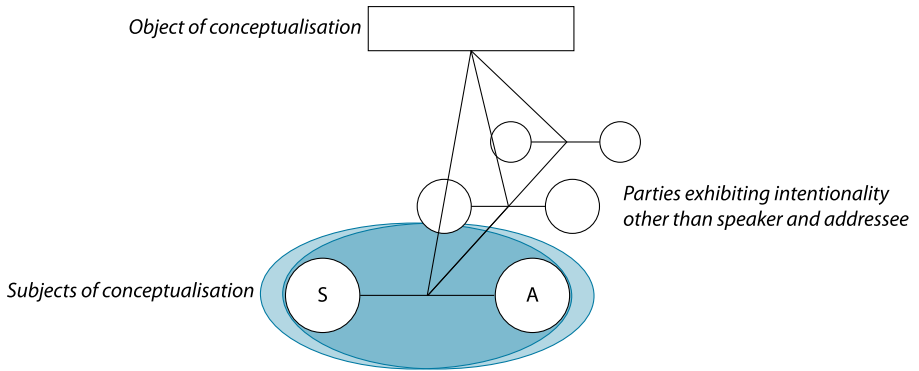


Figure 6. While communicating about some object of joint attention, S and A may refer to other interaction events, each featuring *their* participants

In the situation depicted in Figure 6, viewpoints from the other interaction events must, in one way or another, be relevant to how S and A assess their object of joint conceptualisation. As an example, imagine two people, Simon and Arran, waiting for a man named John to show up at their appointment. Simon has been at John’s house the day before, and when the appointment was mentioned there, John’s daughter Mary brought up that her father is *always* late. In response to that, John pointedly told Mary that he would make sure to be on time on this occasion. Now Simon says to Arran that *John assured Mary that he would be on time*. Figure 7 depicts this situation schematically:

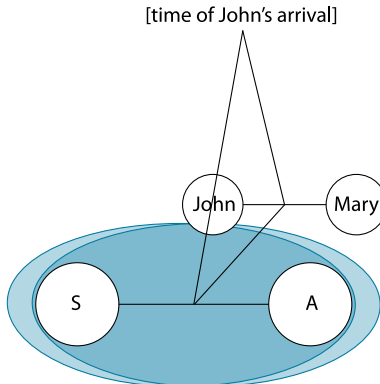


Figure 7. Schematic depiction of *John assured Mary that he would be on time*

In this example, there is thus another interaction event being called up within the current interaction: Simon coordinates his perspective on John’s expected time of arrival with Arran by referring to how John was coordinating *his* perspective on his arrival time with Mary the day before. He could have done this in numerous

alternative ways, for example, by saying *John said to Mary: "I'll make sure to be on time"*, *When I saw John and Mary, John thought he would be on time, John will be on time. He promised Mary*, and so on. All these alternatives feature a slightly different distribution of responsibility for what John said and the amount of commitment made by Simon to John being on time, given that some elements modify the nature of the relationship between third-person conceptualisers and the speaker and addressee in different ways. Thus, in this example, the use of indirect discourse and the choice, by the speaker, of the verb *assure* (unlikely to have been used by John himself), indicate some degree of co-responsibility of (and interpretation of John's utterance by) the speaker, higher than with the use of a neutral verb of speaking and direct discourse (e.g. *John said to Mary: "I'll make sure to be on time"*). These differences are as such interesting from a semantic, grammatical, or narratological perspective, but go beyond the point we want to make here – which is that all alternatives feature different linguistic elements (words, grammatical constructions, patterns of speech and thought representation) with different meanings, leading to a variety of overall interpretations, by operating on parameters within the same conceptual space: the relationship between the speaker and the addressee ((*x*)-axis), the relationship of the communicative interaction with other interactions featuring third parties ((*z*)-axis), and all of their perspectives on the object of conceptualisation ((*y*)-axis).

Finally, it is worth noting that the conceptual space of intersubjectivity proposed here can accommodate viewpoint configurations comprising conjunct, causally related, or otherwise linked third-party perspectives without further increases in dimensionality (see Van Duijn & Verhagen 2019, for more discussion).

3. Thoughtscapes and the 3D-model: Three case studies

We will now discuss three example cases that come from quite different contexts of linguistic analysis and concern phenomena across traditional syntactic and part-of-speech boundaries: constructions for expressing reported speech, adverbs such as *allegedly* and *accidentally*, and temporal markers involved in viewpoint construction. We will show that all examples revolve around linguistic items that can be considered conventional tools operating along the three axes of our conceptual model.

3.1 Case 1: Murder or accident?

The first case is based on news reports concerning the “Pistorius case”, the tragic shooting of Reeva Steenkamp by athlete Oscar Pistorius that took place on 14 February 2013.¹¹ The question that immediately perturbed everyone was whether it was murder or a terrible accident. The difference between the competing versions of what happened during the night of the shooting are completely dependent on the construal of Pistorius’ intentional state at the moment of pulling the trigger: did he *think* he was shooting at a burglar or did he *know* his girlfriend was behind the bathroom door? The news media not only reported the perspective of the athlete, but also of police detectives, spokespeople, journalists, witnesses, family members, and so on. The result was what we have elsewhere termed a *thoughtscape* (Van Duijn 2018, Van Duijn & Verhagen 2019), a series of perspectives that are mutually connected and embedded in various ways. What could be found in the news reports was what we referred to as *polyphonic* discourse representing this thoughtscape: all kinds of linguistic elements were doing some part of the labour of coordinating the involved perspectives, including grammatical constructions (such as complementation and inquit-constructions), various patterns of reported speech and thought, lexical items (such as *allegedly* and *accidentally*), tense, modality, and more. One of the examples was the following opening sentence from a South-African press release:

- (4) *Athlete Oscar Pistorius **allegedly accidentally** shot dead his girlfriend at his house in Pretoria on Thursday morning, Beeld.com reported.*
(SAPA, ‘Oscar Pistorius shoots girlfriend: report’, 14 February 2013)

As a whole, (4) fits a particular embedding pattern, which we have elsewhere termed an *inquit-construction* (Van Duijn & Verhagen 2019), in which the reported clause precedes the reporting clause (underlined). The inquit-construction does part of the viewpoint coordination work: it attributes the claim that *Pistorius allegedly accidentally shot dead his girlfriend* to the perspective of the newspaper *Beeld.com*. However, there are more viewpoints being coordinated. It is implied by the adverbs *allegedly* and *accidentally* (boldface) that some external source *claims* that Pistorius *did not intend* to shoot his girlfriend. In other words, already on the basis of one sentence, readers are confronted with a thoughtscape

11. Background to the case: South-African athlete Oscar Pistorius shot and killed his girlfriend Reeva Steenkamp on February 14th, 2013. Pistorius is a sports icon also known as ‘the Blade Runner’; his legs were amputated and yet he became a sprinter using carbon-blade prosthetic legs. In the aftermath of the killing, news media have frequently reported details of the court case, police investigations, the personal life of Pistorius and Steenkamp, etc. See Van Duijn 2016 (Ch. 4) for a detailed discussion.

involving three viewpoints, without even counting the perspective of the speaker (i.e. the journalist who wrote the sentence).

We will first concentrate on the reported content of the inquit-construction:

(5) *Pistorius allegedly accidentally shot dead his girlfriend.*

The words *allegedly* and *accidentally* are instantiations of *viewpoint packages* (Van Duijn 2016), words implying a topology of viewpoints, introducing one or several viewpoint layers. In the case of *accidentally*, it is given in this topology that an agent did not intend X, while it is known that the outcome is X. In actual usage this topology is assimilated with details provided in the immediate context. For example, readers of (5) will blend their knowledge of the topology of *accidentally* with “Athlete Oscar Pistorius” and “shot dead his girlfriend”, and take it that he shot her dead, but *did not intend* to do so. In this way, the speaker, by using *accidentally*, invites her addressee to consider the perspective of a third party, in this case Pistorius. Schematically:¹²

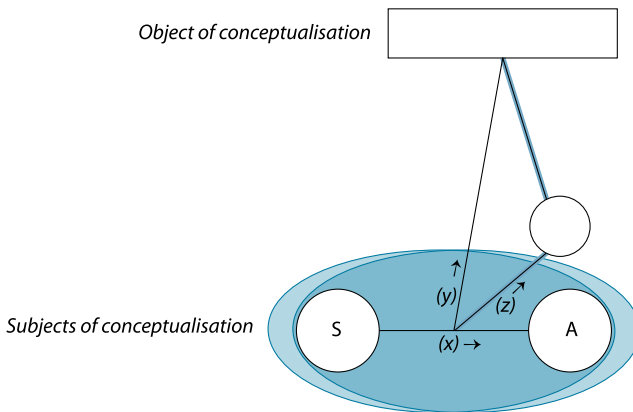


Figure 8. Schematic depiction of *accidentally* in sentence (5) using our 3D conceptual model. Since the perspective of the third-party discourse participant is not highlighted explicitly, the (z)-axis is marked using a light blue line

In a similar way, in the topology of *allegedly* it is given that some source X, not the speaker, asserts the content under the scope of this adverb. This topology can be elaborated to various degrees. The identity of source X can be given in the context, or left open, as is the case in (5): readers of this sentence will understand that some

12. The word *accidentally* clearly also negotiates a relationship to an object in the world on the (y)-axis, but in our discussion here we will abstract from these relationships and focus on those indicated on the (x)- and (z)-axes.

external source not specified here claims that Pistorius accidentally shot dead his girlfriend. In that sense *allegedly* shows strong similarities to the Dalabon element *yangdjehneng* cited in (1) above. It suggests the presence of an extra viewpoint, lowers the epistemic commitment the speaker makes to the related content, and, indeed, could also be ‘translated’ using a complementation construction:

(6) *It is claimed that Pistorius accidentally shot dead his girlfriend.*

In terms of the present model, *allegedly* is thus a linguistic cue that negotiates a particular epistemic stance of the speaker, while at the same time inviting the addressee to consider the perspective of a third, in this case unspecified, party. It operates along the (x)- and (z)-axes, albeit without a particular emphasis on either. Consider the schematic depiction in Figure 9:

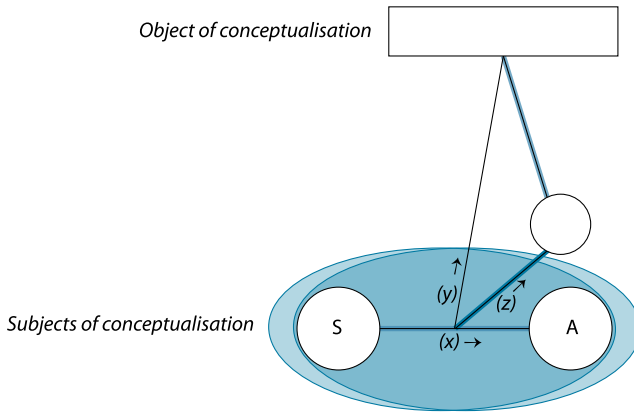


Figure 9. Schematic depiction of *allegedly* in sentence (5) using our 3D conceptual model. The adverb *allegedly* operates implicitly on both the (x)- and (z)-axes, both marked with a light blue line

A schematic rendering of sentence (5), involving at the same time the viewpoint coordination effected by *accidentally* (Pistorius *not intending* to shoot his girlfriend) and *allegedly* (the speaker relegating responsibility for this information to a third, non-specified party), is also possible using the proposed conceptual model and can be depicted as follows (Figure 10):

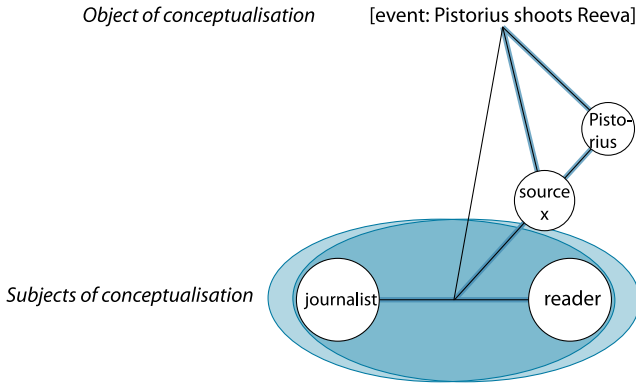


Figure 10. Schematic depiction of viewpoint coordination in sentence (5). A second viewpoint is now included along the (z)-axis

Finally, the perspective of *Beeld.com*, which is coordinated with respect to the reported content using the inquit-construction in (2), can be added to the picture (see Figure 11):

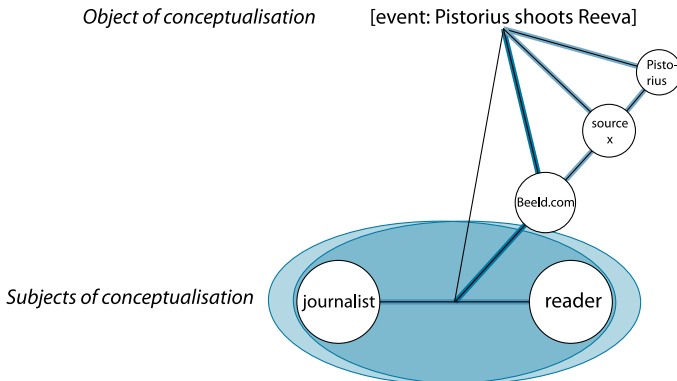


Figure 11. Schematic depiction of sentence (4). Given that the introduction of the perspective of *Beeld.com* takes place explicitly, ‘on stage’, a dark blue line is used here along the (z)-axis. The introduction of the two additional perspectives (Source X and Pistorius) as well as the negotiation of epistemic commitment is done implicitly, ‘off stage’, hence the light blue lines

3.2 Case 2: Scope and framing in adjectives

In his 2002 paper on structure and ordering of pronominal adjectives in English, Vandelanotte addresses adjectives that involve “‘interpersonal’ rather than

‘dependency’ relationships” (Vandelanotte 2002:245), that is: they are not involved in instantiation (e.g. *high* in a *high eighteenth-century façade*) or type specification (*eighteenth-century*) but instead operate on the level of signalling a particular stance from speaker to addressee. Moreover, within this subcategory of interpersonal adjectives he applies the distinction between scope and framing proposed by McGregor (1997): adjectives “expressing strong speaker’s attitudes involve scope”, those “which set something apart as originating in someone else’s discourse, involve framing” (Vandelanotte 2002:246).¹³ As an illustration, consider the following two diagrams, based on those used by Vandelanotte and McGregor:

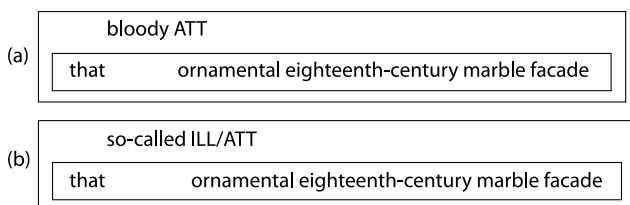


Figure 12. Placement of “bloody” and “so-called” in the outer boxes indicates that these adjectives have scope over the entire noun phrases

The role of *bloody* in Figure 12(a) is indeed not to restrict, modify, or qualify the denotation of the head noun, but to express an attitude of the speaker with respect to the referent at the specific moment of interaction – in this case probably annoyance (...*that bloody ornamental eighteenth-century marble façade is blocking my view!*). In Vandelanotte’s (2002) analysis, the attitude (marked ATT) signalled by *bloody* has scope over the entire noun phrase (as indicated by its position in the outer box). Similarly, *so-called* in Figure 12(b) does not add a property in the way that the adjectives to its right do, but it sets the entire noun phrase apart (again indicated by its position outside the inner box). Vandelanotte points out that the difference between the two consists in *so-called* actually having a twofold function. First, the speaker makes clear that the designation that follows is not his, but that someone else is to be held responsible (this is what McGregor calls its illocutionary force). Second, he signals some of his own opinion of the referent: the façade involved might in fact not look that ornamental, eighteenth-century, and marble at all in the eyes of the speaker (just as a *so-called expert* or an *alleged expert* may in the end not have that much expertise, to para-

13. It should be noted that “framing” in “framing adjective” (e.g. McGregor 1997: 66–67) goes back to a different tradition of use than the one represented by, for example, Fauconnier (1994).

phrase Vandelanotte's 2002:249 discussion). In contrast, *bloody* in Figure 12(a) only has a role of this latter, attitudinal kind.

Thus Vandelanotte goes beyond the traditional distinction between adjectives involved in type specification and those "representing instead an editorial comment by the speaker" (Langacker 1991:59). He works out a useful subdivision within the latter category between scopal adjectives, negotiating speaker's attitudes, and framing adjectives, capable of "set[ting] something apart as belonging to [...] the reality of another's discourse" (Vandelanotte 2007:360). While the distinction is clearly adequate and useful, it must be noted that Vandelanotte has to use special descriptive tools to distinguish them, expressed in the labels ILL and ATT in Figure 12. The model we proposed above, however, already provides the necessary distinctions, and also a basis for explaining why ILL always implies some ATT, so to speak. Besides the reference to another viewpoint, *so-called* and *alleged* also operate on the level of stance negotiation between speaker and addressee. All instances mentioned here lower the commitment the speaker exhibits to the truth or validity of the content in the scope of the adjective or adverb. In some cases there is also some signalling of negative attitude: for Figure 12(b) it seems fair to conclude that the speaker in fact has a lower opinion of the object of attention than the party whose viewpoint is evoked. So, in this respect, the role of framing adjectives resembles that of *bloody* from Figure 12(a). In our model, both this epistemic stance and signalling of negative attitude can be indicated along the (*x*)-axis. What sets *so-called* and *alleged* (and other framing adjectives) apart from scopal adjectives such as *bloody* is that they operate along the (*z*)-axis: they invite the addressee to consider another party's perspective on the object of joint attention. But as we said when we introduced the model, a third party's perspective always *affects* the speaker's and addressee's negotiation of how the common ground should be updated with respect to their object of joint attention; in other words, third-party perspectives are ultimately always *instrumental* in the process of alignment between speaker and addressee in the (common) ground. This is, in our view, precisely the reason why there are no *framing* adjectives, highlighting the (*z*)-axis between the ground and a subject of conceptualisation external to it, without some attitudinal consequence on the (*x*)-axis, while there are *scopal* adjectives like *bloody* with an attitudinal role, but no additional viewpoint layer.

By drawing attention to this, we have incorporated the subdivisions in the domain of attitudinal adjectives made by Vandelanotte into the arena of the conceptual space we proposed to model linguistic interaction more generally. The widely observable phenomenon of discourse exhibiting a polyphony of intentional states is simply also manifest in the domain of adjectives: some only coordinate the intentional states of speaker and addressee, some also involve other

minds in this coordination process. The difference is depicted schematically in Figures 13 and 14.

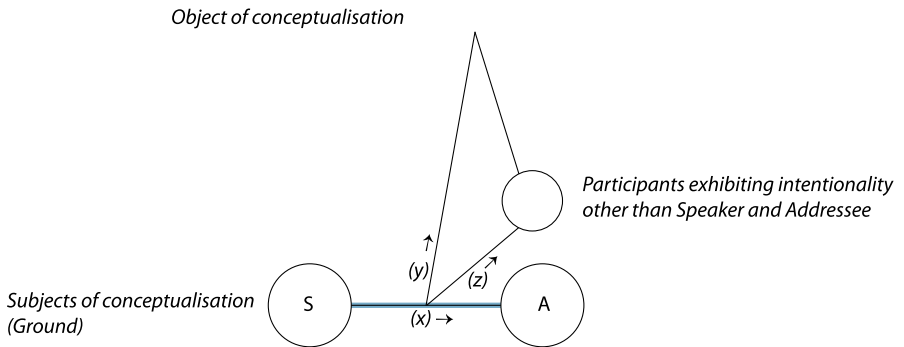


Figure 13. Schematic depiction of *bloody* as used in the example from Vandelanotte (2002: 245). The blue line along the (x)-axis indicates that this adjective coordinates the perspectives of the Signaller and the Addressee

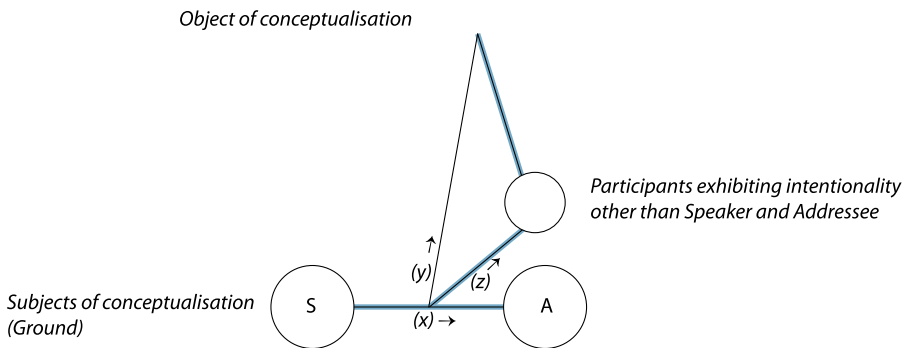


Figure 14. Schematic depiction of *so-called* as used in the example from Vandelanotte (2002: 246). The blue lines along the (x)- and (z)-axes indicate that this adjective coordinates the perspectives of the Signaller, the Addressee, and an unspecified third party. The same figure could be used to schematically represent *alleged*

Other examples of framing adjectives besides *so-called* and *alleged* are *reported*, *purported* and *supposed*. These all have in common that they relegate responsibility for parts of the presented content from the speaker to another conceptualiser. Or in terms of Van Duijn (2016: Ch. 4): they are viewpoint packages bringing in another viewpoint layer as part of their topology. It is possible to elaborate this topology by indicating in the context who is responsible for the extra viewpoint, or to leave this open. As the examples below suggest for the adverb *allegedly*, this elaboration can be done to various degrees:

- (7) *one tries to sell [...] smear in a bottle allegedly from the great Madonna herself* (BNC, CBC)
- (8) *Simon Peter [...] on whom Jesus allegedly founds his church* (BNC, EDY)
- (9) *the protesters' vociferously expressed and allegedly 'sincere' ideals* (BNC, HTP)

Whereas in (7) the source of viewpoint referred to using *allegedly* is some unknown (but likely questionable) authority regarding the authenticity of a pop-music relic, in (8) it is quite clear from the context that the source referred to is the Christian tradition, but no particular individual is specified. This is different in (9): here it is clear that the party responsible for the designation of the ideals as '*sincere*' is formed by the protestors mentioned earlier.¹⁴ All of these options are inherently available in the three-dimensional model we are proposing.

3.3 Case 3: Perspective mixing: Free Indirect Discourse, and other forms

Finally, in this section we will show how the 3D-model can incorporate an account of some uses of Free Indirect Discourse and other forms of polyphony in a natural way. As a start, consider the following small fragment from the novella *The Tiger Moth* by H. E. Bates (Westh 1980, cited in Hutchinson 1988):

- (10) *She rested her fork on the edge of her plate and he noticed for the first time that she was wearing no wedding ring. He immediately changed the subject.*

The first linguistic viewpoint indication here is the complement-taking verb *notice* embedding the indirect thought *that she was wearing no wedding ring*. The meaning of this verb – a kind of perception verb – marks the male character as the focaliser in this passage (Genette 1980, Bal 1985). But there is more than only 'seeing' going on; readers understand that they are also presented with some internal thinking process of this character. In fact, this *must* be the case, as the *absence* of something (here a wedding ring) cannot, as a matter of principle, be *perceived*. The absence of a wedding ring is an *inference* based on some actual perception (say, of the woman's hand) and a background assumption (say, a woman of this age may be expected to be married), hence a conceptual rather than a perceptual operation.¹⁵

14. Note the quotation marks in (9): it is unclear from a sample of 100 instances in the British National Corpus we have looked at whether the use of quotation marks is indicative of this type of use.

15. More generally, the use of negation (*no*) can be described in terms of setting up a counterfactual "mental space" (Fauconnier 1994: 32, 96–98, Verhagen 2005: 29–35) in which the positive

In terms of our model, then, what we read in (10) is a reasoning process that is *primarily* located in the male character, not in (one of) the participants in the communicative situation. The mental states of the latter are being coordinated via a connection to another perspective that is marked with the factive perception verb *noticed*, thus implying the third dimension of the (z)-axis. Schematically:

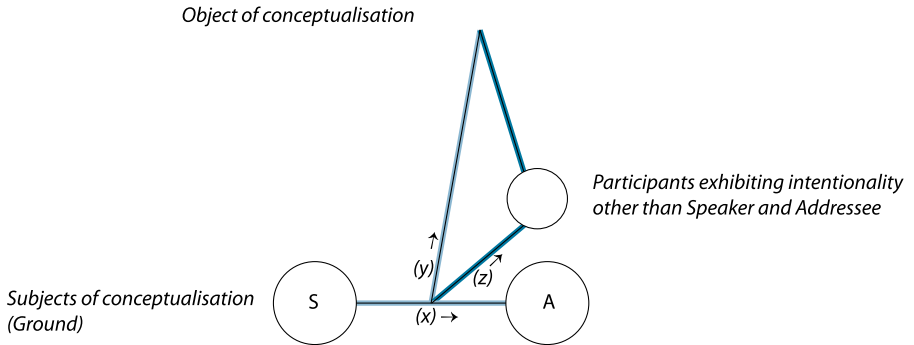


Figure 15. Schematic depiction of fragment (10), where Signaller and Addressee coordinate their perspectives with respect to an object of conceptualisation via the implied reasoning process located in a third party

In this analysis, the narrator also informs the reader that the female character is not wearing a wedding ring, i.e. evoking-and-contradicting a natural expectation, but now indirectly – hence the blue lines – mediated by the male character’s perceptual and conceptual processes. The conceptual structure evoked by the negation element itself is not different from that in the case of Figure 14; its primary application is just ‘shifted’ one layer of viewpoints away from the ground. We have seen this kind of shifting operating before, especially in our discussion of the thoughtscape involving the elements *allegedly*, *accidentally*, and *Beeld.com reported* in sentence (4) (Section 4.1).

Perspectivising elements are thus capable of inducing shifts along the (z)-axis when contextually appropriate ‘anchors’ are found without changing their intrinsic functional role. This idea turns out to be useful in other respects as well. Consider fragment (11), which includes (10) and the passage that immediately follows it:

- (11) *She rested her fork on the edge of her plate and he noticed for the first time that she was wearing no wedding ring. He immediately changed the subject. ‘Are you in one of the services?’ he said.*

proposition holds, and which is contradicted from the “reality space” (Fauconnier) or ground (Verhagen).

No, she said, she was teaching literature and history in St. Anne's High School for Girls. They had been evacuated from London to a mansion called Clifton Court. Did he know it?

'I see it from the air. Sounds pretty dull though. Still, fun and gossip in the common room I've no doubt.'

No, she was free of all that, she said, thank God. She'd managed to buy a small cottage of her own.

'Sounds cosy. Perhaps I might invite myself over some time?'

'The garden's a mass of weeds.'

This enigmatic answer of hers had the effect of changing his interest into a certain excitement.

We can observe an alternation of Direct and Free Indirect Discourse (FID) here. Up to a point, *his* turns are represented directly, whereas *hers* are formulated as FID, in a highly significant way (as noted in Hutchinson 1988). Interestingly, though, the traditional narratological and stylistic definition of FID as mixing of the narrator's and a character's voices (cf. Leech & Short 2007) does not fit this case. The point is that the presentation of her utterances is not directly mediated by the narrator, but by the male character, construed as the focaliser at the beginning: we hear her turns *as perceived and understood by him* – i.e. it is a case of mixing of voices of two *characters*. The ultimate effect is that readers understand that it is easy for him to integrate her contributions into his understanding of the developing conversation, up to the point where she says *The garden's a mass of weeds*: now we hear her utterance directly, no longer mediated by his understanding, which invites the inference that such understanding is lacking (cf. the subsequent characterisation *enigmatic*).

In the present framework, this phenomenon can be easily accommodated, with no need for terminological differentiation or exception conditions. Like Vandelanotte's terminological distinction in the previous section, the difference between mixing narrator's and character's voices on the one hand, and mixing different characters' voices on the other, can be made to follow from the presence of the (z)-axis in our model. The grammatical and lexical viewpoint tools (tense, deictic reference, negation, interrogatives, etc.) all preserve their intrinsic semantic-structural properties, but they may be used in managing relations between any two viewpoints along this axis: in the most straightforward cases between the ground and one layer 'down', but also between other layers (ultimately always relating back to the ground). We hypothesise that upon close inspection this phenomenon will be found to occur quite regularly, and that it may often go unnoticed precisely because of the preservation of the intrinsic properties of the linguistic elements involved.

The 3D-model thus invites a view of linguistic viewpoint-indicating tools as relatively free to combine, as long as the resulting thoughtscape makes sense for the overall discourse. In fact, this precisely appears to be the position that we need in order to accommodate several observations in the literature suggesting that the traditional tripartite distinction between Direct Discourse (Speech or Thought), Indirect Discourse, and Free Indirect Discourse does not exhaust the possibilities of mixing. One example is Vandelanotte's (2004, 2012) observation about possible patterns of coreference between independent reported clauses and postposed reporting clauses:

- (12) *I will be late, John said* [Direct Discourse]
 (13) *He would be late, John said* [Free Indirect Discourse]
 (14) a. *John will be late, he said* (in a conversation) [??]
 b. *John would be late, he said* (in a story)

Patterns like (14), with *John* and *he* understood as coreferential, exhibit a mixed form, like FID, but they differ even more from Direct Discourse (in English) in that the choice of a proper name for the subject in the reported clause is more of a narrator's intervention than the shift from first to third person pronoun between (12) and (13) – in our culture, speakers do not normally refer to themselves with their name. Vandelanotte therefore proposes yet another category of mixed discourse representation: Distanced Indirect Speech-or-Thought (DIST). His characterisation of this as involving a greater degree of narrator responsibility and less for the character than FID seems quite to the point, also in view of the interpretation of a Dutch internet example like (15), which indeed suggests that the reported clause may be a kind of summary, and not at all a report that allows an accurate reconstruction of the actual speaker's words.

- (15) *Een paar uur [eerder] was ploeggenoot Wout Poels uit de Tour de France gestapt. Voor een camera van de NOS legde hij uit waarom. Wout was ziek, zei hij. Voor de rest heb ik niet veel kunnen volgen van hetgeen Wout vertelde.*
 'A few hours [earlier], team mate Wout Poels had abandoned the Tour de France. In front of an NOS camera, he explained why.
 Wout was ill, he said. Otherwise I could not get much of what Wout was telling.'

Given our 3D-model, this interpretation and especially the difference between patterns of the type in (13) and those of the type in (14) can also be characterised directly in terms of the intrinsic function of proper names on the one hand, and

the fact that they are used in a reported clause on the other – again without the need for a proliferation of categories of Speech-and-Thought representation.

Another case in point is the idea that the representation of Speech should be categorically kept distinct from that of Thought, because the norm for representing the former is Direct and that for the latter is Indirect (only speech being observable; cf. Leech & Short 2007: 276). Now consider the following fragment, discussed in Sanders (2010; comment between parenthesis in her original):

(16) *De huisarts, die op verzoek van Carla met spoed langs komt, ziet een voldragen baby, een meisje van ruim zeven pond. Later wordt vastgesteld dat het kind weken tevoren is overleden, maar wel heeft geleefd.*

Carla denkt dat Etta de moeder is. Van wie anders zou het kindje moeten zijn?

‘The family doctor, who at Carla’s request arrives within moments, sees a full-term infant, a girl well over seven pounds. Later it is determined that the child died weeks before, but did live.

Carla thinks that Etta is the mother. Who else could the baby belong to? (*lit.* Of who else should the baby – *lit.* the child + diminutive suffix – be?)’

The question is: what category does the sentence starting with *Carla thinks...* belong to? One might say: thought representation, because of the verb *think*. However, the preceding context makes it likely (though it is not even hinted at) that Carla and the doctor have had a conversation – after all, it would be quite unnatural for the doctor to have come over at Carla’s request, and that the two of them then do *not* talk about the rather gruesome discoveries and what might be behind them. The rhetorical question at the end would also fit better in such a conversation frame than as only a self-addressed one. But in fact, the previous sentence also allows for an interpretation as reported *speech*, as a consequence of a rather general property of complementation constructions. Consider (17):

- (17) *John promised that he would be on time.*
 a. *John said: “I promise that I will be on time”*
 b. *John said: “I will be on time”/...*

This may be a truthful report of John himself having used the verb *promise* (cf. 17a), in which case he is the one primarily responsible for that characterisation, but also of John having used rather different words (cf. 17b), in which case it is the present speaker who is to be held responsible for reporting them as a promise. This is a general option for the interpretation of the matrix verbs in instances of the complementation construction (Verhagen 2005: ch. 3),¹⁶ not only for verbs of

16. There are some specific verbs that block this kind of interpretation, for instance *threaten* (cf. Verhagen 2000).

communication – it also holds for *think* (and for the verb *notice* in (10)). So *Carla thinks that Etta is the mother* may well be read as a report of Carla having said: “*I think that Etta is the mother. (Who else could the baby belong to?)*”, thus a case of reported speech. As observed before, this would fit the context really well too, but more important for our present purposes is that the generality of this option, in languages that have complementation-like constructions, undermines a categorical distinction between speech and thought representation (and after all, the best evidence for reporting what a third person thinks is to refer to what that person has explicitly stated...). Rather, in terms of the 3D-model, we should simply say that complementation constructions establish a characterisation along the (z)-axis of a third party’s mental state for which the present speaker takes at least some responsibility (cf. Verhagen 2005:104–119), leaving the amount of responsibility of the third party open in principle (with some lexical items imposing specific constraints; cf. footnote 16).

As a final example of this kind of flexibility in combinatorial options, consider the following caption placed below a photograph in the Dutch newspaper *Spits* of August 23rd, 2011:

- (18) *Kamermeisje Nafissatou Diallo (links) en haar advocaat spreken gisteren met de pers*
 ‘Chambermaid Nafissatou Diallo (left) and her lawyer talk to the press yesterday’

The photograph shows Diallo standing next to her lawyer, who is clearly speaking to an audience in front of them. Formally, the combination of present tense (*talk*) and the deictic adverb *yesterday* in the caption is the exact mirror image of the pattern considered typical for FID (cf. Nikiforidou 2010 on “was-now”). But it is still a perfectly felicitous combination. While the past tense in prototypical FID fits the story telling and the temporal adverb the actuality for the character, the present tense of the action (of talking) in (18) fits the immediate presence of the picture and the temporal adverb the relative time of the event. In other words, when each of the deictic elements can find an ‘anchor’ in its context – however that is provided, by the co-text in a story, by a picture in a newspaper, etc. – in such a way that the overall configuration in the three-dimensional model of intersubjectivity makes sense to a reader for updating the common ground, then this combination can be communicatively successful.

4. Conclusion and discussion

4.1 Summary

All in all, using the conceptual framework suggested here, and depicted schematically in several figures above, we argue that linguistic elements across different categories, levels of analysis, and languages (lexical units, grammatical and narratological patterns, English, Dalabon) operate along three dimensions: speaker and addressee negotiate ((x)-axis) how the common ground should be updated with respect to a particular object of conceptualisation ((y)-axis), potentially by inviting the other to view this object of conceptualisation (in part) from the perspective of third parties ((z)-axis). In the case of some interaction events this process of updating the common ground involves no third-party perspectives at all (to those interaction events only the first two dimensions are relevant), whereas in other cases a complex thoughtscape can be conjured up in the course of this process. Sometimes the perspectives in these thoughtsapes are embedded (cf. Figures 10 and 11) and sometimes they are related in different ways. For example, third parties can themselves be represented as being interlocutors in a different interaction event (cf. Figure 7).

4.2 The broader context: Intersubjective language and intersubjective thought

Finally, widening the perspective, we would like to point out a potential way of connecting the three-dimensional nature of intersubjective communication as proposed here to a fundamental feature of human interaction as discussed in the literature. It has been widely argued that humans are exceptionally good *mindreaders* (also often referred to as *theory of mind*, see e.g. Apperly 2011 and Van Duijn 2016 for overviews and discussions). In many situations of daily life we form understandings of the inner lives of others, and use these for making sense of the social world around us. Although theories vary, researchers generally agree that the special, adult-human mindreading competence must in part be explained by genetically rooted factors exerting their force directly from birth, and partly by factors in the socio-cultural environment children grow up in (Wellmann 2014). In our view, mindreading is best seen as a case of “dual inheritance” (McElreath & Henrich 2007), i.e. its present form has been produced by a combination of processes of genetic and cultural evolution. Children who grow up in our species-specific socio-cultural environments clearly have ample opportunity to observe social interaction and practice it themselves, but they also benefit greatly from the acquisition of language. Not only does it open up the possi-

bility of receiving explicit instructions, corrections, and explanations from caretakers and peers, but languages also form culturally evolved ‘storage devices’. Over generations they have accumulated information about how members of a cultural-linguistic community have found it useful to represent the world and divide it into categories, in order to be able to communicate in effective and efficient ways. Learning and internalising the rules of use of their meaningful elements therefore help in building the scaffolds for useful strands of reasoning, not replacing but augmenting existing (innate) capabilities (cf. Tomasello 2014, Heyes 2018, Van Duijn 2016).

A subset of the meaningful elements of every language consists of those in focus in this paper: linguistic items capable of highlighting and/or mutually coordinating the perspectives of ourselves, our interlocutors, and third parties. We would expect that acquiring these linguistic tools supports children’s understanding of the working of other minds, and thus their mindreading abilities: not only do they gain access to a realm of information about how others view the world, they also internalise the *structural* characteristics of human intersubjective interaction. This fits with evidence from various studies investigating language acquisition in relation to mindreading (see Milligan et al. 2007 for an overview; cf. also Berger & Luckmann 1966).

The structure of our model, and the way it shows a degree of complexity beyond that of much animal communication, does not necessarily correspond to the actual path of the evolution of communication in our species, or, for that matter, to the development of communicative abilities in children – each of these and similar issues requires independent empirical and theoretical research. But a proper conceptualisation of the structure of human communication does inform and constrain such research, as it specifies what it is that has evolved and develops, and it is also to this end that we have attempted in this paper to show and justify here that such a conceptualisation has to go beyond triadic, ‘two-dimensional’ communication.

References

- Apperly, Ian. 2011. *Mindreaders: The cognitive basis for theory of mind*. New York: Psychology Press.
- Bakhtin, Mikhail M. 1981. *The dialogic imagination*. Edited by Michael Holquist, translated by Caryl Emerson and Michael Holquist. Austin/London: Texas University Press.
- Bal, Mieke. 1985. *Narratology: Introduction to the theory of narrative*. Toronto: Toronto University Press.
- Berger, Peter L. & Thomas Luckmann. 1966. *The social construction of reality: A treatise in the sociology of knowledge*. Harmondsworth: Penguin.

- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511620539>
- Dancygier, Barbara. 2012. *The language of stories: A cognitive approach*. New York: Cambridge University Press.
- Dancygier, Barbara, Wei-lun Lu & Arie Verhagen (eds.). 2016. *Viewpoint and the fabric of meaning: Form and use of viewpoint tools across languages and modalities*. (Cognitive Linguistics Research [CLR] 55). Berlin/Boston: de Gruyter Mouton.
<https://doi.org/10.1515/9783110365467>
- Dennett, Daniel C. 1987. *The intentional stance*. Cambridge MA: MIT Press.
- Du Bois, John W. 1985. Competing motivations. In John Haiman (ed.), *Iconicity in syntax*, 343–365. Amsterdam: Benjamins. <https://doi.org/10.1075/tsl.6.17dub>
- Du Bois, John W. 2003. Discourse and grammar. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, vol. 2. 47–87. London: Erlbaum.
- Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139–182. Amsterdam: Benjamins.
<https://doi.org/10.1075/pbns.164.07du>
- Duijn, Max J. van. 2016. *The lazy mindreader: A humanities perspective on mindreading and multiple-order intentionality*. PhD Thesis, Leiden University.
- Duijn, Max J. van. 2018. Readers' mindreading challenges, and how they can inform cognitive science. *Review of General Psychology* 22(2). 188–198.
- Duijn, Max J. van, & Arie Verhagen. 2019. Recursive embedding of viewpoints, irregularity, and the role for a flexible framework. *Pragmatics* 29(2). 198–225.
- Evans, Nicholas. 2010. *Dying words: Endangered languages and what they have to tell us*. Chichester: Wiley-Blackwell.
- Fauconnier, Gilles. 1994. *Mental Spaces: Aspects of meaning construction in natural language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511624582>
- Fludernik, Monika. 1993. *The fiction of language and the languages of fiction*. London: Routledge.
- Genette, Gérard. 1980. *Narrative discourse: An essay in method*. Ithaca, NY: Cornell University Press.
- Heyes, Cecilia. 2018. *Cognitive gadgets: The cultural evolution of thinking*. Cambridge, MA: Harvard University Press.
- Hühn, Peter, Wolf Schmid & Jörg Schönert (eds.). 2009. *Point of view, perspective, and focalization: Modeling mediation in narrative*. Berlin: Walter de Gruyter.
<https://doi.org/10.1515/9783110218916>
- Hutchinson, Tom. 1988. Speech presentations in fiction with reference to *The Tiger Moth* by H. E. Bates. In Michael H. Short (ed.), *Reading, analysing and teaching literature*, 120–145. London: Longman.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Volume I: Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar: Volume II: Descriptive application*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2014. Subordination in a dynamic account of grammar. In Laura Visapäa, Jyrki Kalliokoski, Helena Sorva (eds.). *Contexts of subordination: Cognitive, typological and discourse perspectives*, 17–72. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Leech, Geoffrey & Mick Short. 2007 [1981]. *Style in fiction: A linguistic introduction to English fictional prose*. 2nd edn. Harlow: Pearson Education Limited.
- McElreath, Richard & Joseph Henrich. 2007. Dual inheritance theory: The evolution of human cultural capacities and cultural evolution. In Robin Dunbar & Louise Barrett (eds.), *Oxford handbook of evolutionary psychology*. Oxford: Oxford University Press.
- McGregor, William B. 1997. *Semiotic grammar*. Oxford/New York: Oxford University Press.
- Milligan, Karen, Janet Wilde Astington & Lisa Ain Dack. 2007. Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development* 78(2). 622–646. <https://doi.org/10.1111/j.1467-8624.2007.01018.x>
- Nikiforidou, Kiki. 2010. Viewpoint and construction grammar: The case of *past+now*. *Language and Literature* 19(3). 265–284. <https://doi.org/10.1177/0963947010370253>
- Owings, Donald H. & Eugene S. Morton. 1998. *Animal vocal communication: A new approach*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139167901>
- Sanders, José. 2010. Intertwined voices: Journalists' modes of representing source information in journalistic subgenres. *English Text Construction* 3(2). 226–249. <https://doi.org/10.1075/etc.3.2.06san>
- Sanders, Ted, José Sanders & Eve Sweetser. 2009. Causality, cognition and communication: A mental space analysis of subjectivity in causal connectives. In Ted Sanders & Eve Sweetser (eds.), *Causal categories in discourse and cognition*, 19–59. Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110224429.19>
- Seyfarth, Robert M., Dorothy L. Cheney & Peter Marler. 1980. Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour* 28(4). 1070–1094. [https://doi.org/10.1016/S0003-3472\(80\)80097-2](https://doi.org/10.1016/S0003-3472(80)80097-2)
- Sweetser, Eve. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620904>
- Tomasello, Michael. 2008. *Origins of human communication*. Boston: MIT Press.
- Tomasello, Michael. 2014. *A natural history of human thinking*. Cambridge/London: Harvard University Press. <https://doi.org/10.4159/9780674726369>
- Vandelanotte, Lieven. 2002. Prenominal adjectives in English: Structures and ordering. *Folia Linguistica* XXXVI 3–4. 219–259.
- Vandelanotte, Lieven. 2004. Deixis and grounding in speech and thought representation. *Journal of Pragmatics* 36(3). 489–520.
- Vandelanotte, Lieven. 2007. *Mister so-called X*: Discourse functions and subjectification of *so-called*. In Christopher S. Butler, Raquel Hidalgo Downing & Julia Lavid (eds.), *Functional perspectives on grammar and discourse: In honour of Angela Downing*, 359–394. Amsterdam: Benjamins. <https://doi.org/10.1075/slcs.85.20van>
- Vandelanotte, Lieven. 2009. *Speech and thought representation in English: A cognitive-functional approach*. Berlin/New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110215373>
- Vandelanotte, Lieven. 2012. “Wait till you got started”. How to submerge another’s discourse in your own. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 198–218. Cambridge: Cambridge University Press.
- Verhagen, Arie. 2000. “The girl that promised to become something”: An exploration into diachronic subjectification in Dutch. In Thomas F. Shannon & Johan P. Snapper (eds.), *The Berkeley Conference on Dutch Linguistics 1997: The Dutch Language at the Millennium*, 197–208. Lanham, MD: University Press of America.

- Verhagen, Arie. 2005. *Constructions of intersubjectivity: Discourse, syntax, and cognition*. Oxford: Oxford University Press.
- Verhagen, Arie. 2008a. Intersubjectivity and explanation in linguistics – A reply to Hinzen and Van Lambalgen. *Cognitive Linguistics* 19(1). 125–143. <https://doi.org/10.1515/COG.2008.007>
- Verhagen, Arie. 2008b. Intersubjectivity and the architecture of the language system. In Jordan Zlatev, Timothy P. Racine, Chris Sinha & Esa Itkonen (eds.), *The shared mind: Perspectives on intersubjectivity*, 307–331. Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/celcr.12.17ver>
- Wellman, Henry M. 2014. *Making minds: How theory of mind develops*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199334919.001.0001>
- Westh, Cecil. 1980. *H. E. Bates: The song of the wren, the dam, the tiger moth, Oh! Sweeter than the berry*. Copenhagen: Gyldendal.
- Wilkins, David. 1986. Particle/clitics for criticism and complaint in Mparntwe Arrernte (Aranda). *Journal of Pragmatics* 10(5). 575–596. [https://doi.org/10.1016/0378-2166\(86\)90015-9](https://doi.org/10.1016/0378-2166(86)90015-9)

Address for correspondence

Max van Duijn
 Leiden Institute of Advanced Computer Science
 Niels Bohrweg 1
 Leiden 2333 CA
 The Netherlands
 m.j.van.duijn@liacs.leidenuniv.nl

Biographical notes

Max van Duijn is assistant professor at Leiden University's Institute of Advanced Computer Science (LIACS), where he co-founded the Creative Intelligence Lab. This lab brings together researchers from the cognitive and computer sciences who have a shared interest in the foundations of intelligence, both “in carbo” and “in silico”, and who recognise creativity and playfulness as key factors in scientific innovation. He has a background in linguistics, literature, and cognitive science. His PhD thesis *The Lazy Mindreader* (Leiden and Oxford; defended 2016) studied multiple-order intentionality (A believes that B wants that C thinks...etc.) in relation to language and narrative. In his current research he continues to connect topics and methods from the sciences and humanities, focusing particularly on the links between socio-cognitive development and narrative competence in children. His work is published in English and Dutch, in scholarly as well as popular venues.

Arie Verhagen (PhD 1986) is professor emeritus of Language, Culture, and Cognition at Leiden University (The Netherlands). He previously held positions at the Free University in Amsterdam, the University of Utrecht, Leiden University, and the University of Antwerp. From 1996 to 2004 he served as editor-in-chief of the journal *Cognitive Linguistics*. His grammatical work includes studies on word order, passive, causative, connectives, *wh*-questions, complementa-

tion, and other construction types. With his 2005 monograph *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition* (Oxford University Press), he contributed to the so-called 'social turn' in cognitive linguistics. His research is framed in a (radically) usage-based approach (for an overview, see Dirk Geeraerts, "Grammar in the context of intersubjective usage", *Nederlandse Taalkunde/Dutch Linguistics* 21 (2016), 395–407), and focuses especially on the connection between grammar, discourse, and the highly developed human ability to understand other minds, as a basis for cooperation.

Publication history

Date accepted: 5 February 2019