

RESEARCH

Confronting perspectives: Modeling perspectival complexity in language and cognition

Sonja Zeman

LMU Munich/University of Bamberg, DE

Sonja.Zeman@lmu.de

In search of a deeper understanding of perspective-taking, the paper aims at isolating the basic components of complex perspectivization tasks. Based on the assumption that perspective is a relational concept constituted by the dynamic relation between a conceptualizer's point of view and a conceptualized object and, as such, a concept which is shared by both language and cognition, the paper draws a comparison between the developmental stages of cognitive perspectival capabilities, grammatical perspectivization, and viewpoint comprehension in narrative discourse. This comparison leads to a multi-stage model of degrees of perspectival complexity, whereby the most complex tasks in both language and cognition are characterized by a three-point constellation that integrates contrasting internal vs. external perspectives under a superordinated monitoring viewpoint. As a result, the paper argues that a concept of perspective 'taking' is not sufficient to capture the specific aspects of L(anguage)-perspectivization, but has to be complemented by a concept of 'confronting' perspectives.

Keywords: perspective-taking; viewpoint shift; epistemicity; propositional attitudes; narrative discourse; false belief

1 The basic meaning of 'perspective'

The basic meaning of 'perspective' refers to the insight that in visual perception an object is visible only in certain aspects and cannot be viewed in its entirety without changing the observer's point of view. In the real world, we can see either the inside or outside of a house, so the observer has to move his position if he wants to take another perspective. In an abstract sense, 'perspective' can thus be seen as a two-point-relation between an observer's origo and (the aspects of) an object in focus, whereas the process of perspectivization refers to the selection of viewpoint that effects a restriction of the aspects seen (cf. e.g. Graumann 1989: 96).

In this very general sense, the concept of perspective has been metaphorically transferred to the domain of language: language is perspectival in that various options on the lexical, grammatical, contextual, and conversational level allow for conceptualizing the same extra-linguistic situation in different manners so that every utterance is based on a set of choices by the speaker. As a result, the concept of 'perspective' has been applied to various linguistic phenomena in their different aspects (see for an overview e.g. Klein & von Stutterheim 2002; Verhagen 2007; von Stutterheim & Carroll 2007). Furthermore, recent studies have shown that even "seemingly 'innocent' or 'viewpoint-neutral' lower-level constructions" such as determiners or negation can function as viewpoint markers (e.g. Dancygier 2005; 2012a; b; c; Dancygier and Sweetser 2012; Sweetser 2012; Dancygier & Vandelanotte 2016). This insight in the ubiquity of viewpoints has led to the fact that the investigation of multiple viewpoints and their constellations has become a major research topic.

The chapter is based on these studies by taking for granted the ubiquity of viewpoints and the necessity for investigations of multiple viewpoints and their constellations (Dancygier & Vandelanotte 2016). Yet, it is committed to a slightly different perspective. It is not the aim to study different linguistic phenomena with respect to their specific viewpoint potential in discourse, but rather focus on the core concept of perspective itself. In this respect, the main thesis is that the concept of perspective-taking is not sufficient in order to capture the specific aspects of linguistic and cognitive perspectivization but requires the concept of a meta-perspective that is capable to integrate more than one single point of view at a time. So while in reality the local position of the observer restricts her view to perceiving *either* the inside *or* outside of a house, cognitive and linguistic representations do not only enable us to free ourselves from the *here and now* of the actual perceptual situation but also open up the possibility to take and evaluate an internal and external view simultaneously. As a result, a model of L-perspectivization has to go beyond the concept of perspective-taking by taking into account the integration of *confronting* perspectives.

In order to pursue this line of argumentation, the chapter is organized as follows. Based on the assumption that perspectivization in its abstract sense is a fundamental concept which is shared by both language and cognition, Section 2 argues that its basic mechanisms can be isolated by a comparison of cognitive and linguistic perspectival tasks of different complexity. With respect to C(ognition)-perspectivization, Section 3 shows that the intricacy of Theory of Mind (ToM) tasks does not primarily lie within the capability of switching from one perspective to another, but to evaluate confronting perspectives from an additional, external viewpoint. This leads to a differentiation between the concept of perspective-taking as a binary relation between a viewpoint and an object in focus and the ternary concept of *confronting* perspectives. Such refinement is, as Section 4 argues, also needed in order to model the different degrees of perspectival complexity in language. By a comparison of propositional attitude ascriptions (4.1), epistemic meanings of modal verbs (4.2), and narrative discourse comprehension (4.3) it is shown that these different phenomena share the fact that they integrate both external and internal viewpoints on potentially truth-incompatible perspectives. Section 5 will bring the perspectives together by deriving the commonalities between cognitive and linguistic perspectivization. In sum, the analysis leads to a multi-stage-model of different degrees of perspectival complexity that not only refines the concept of perspectivization but also offers a tertium comparationis for complex viewpoint constellations in language and cognition.

2 The intersection of cognitive and linguistic perspectivization

It is well known that in the perceptual domain, children only gradually come to understand that an object can be perceived in a different way from a different point of view, and it is even later that they can judge *how* an object is seen from a different angle (cf. e.g. Flavell 1992; Michelon & Zacks 2006; Moll & Meltzoff 2011a; Surtees et al. 2013). In tradition of the studies of Flavell (e.g. Flavell et al. 1981; Flavell 1992), the acquisition of visual perspectivization is commonly described by referring to the distinction between Level 1 perspectivization, i.e. the ability to understand that someone else might not see an object that oneself can see, and Level 2 perspectivization, i.e. to understand that “an object simultaneously visible to both the self and the other person may nonetheless give rise to different visual impressions or experiences in the two if their viewing circumstances differ” (Flavell et al. 1981: 1). While Level 1 perspective-taking is evident by the age of 24-months (Moll & Tomasello 2006), Level 2 perspectivization develops later at around four years of age (Flavell et al. 1981).¹ The later stage of visual perspectivization

¹ For further differentiation of these two levels see Moll & Meltzoff 2011 and Section 3.

has been seen in close relationship to the emergence of cognitive non-spatial perspectivization tasks, most prominently referred to in terms of *Theory of Mind* (ToM) (cf. for an overview Farrant et al. 2006). ToM refers to the capability of ascribing mental states to other people's minds and is as such a genuine task of perspective-taking. It is, however, not an unproblematic concept since it is controversial which perspectival capabilities in particular are lying behind the understanding of others' beliefs.² Furthermore, ToM has been challenged by approaches which put emphasis on the fact that understanding the belief of others is just one specific aspect of *intersubjectivity* and has to be seen in context of other intersubjective capabilities like comprehension of emotions, attentional foci, intentions, and social embodied interaction (cf. Gärdenfors 2008; Zlatev et al. 2008). One of the main differences in contrast to classical ToM accounts lies within the fact that, instead of focussing on the 'great divide' at four years, also early capabilities like joint attention are seen as capabilities of perspectivization. ToM is thus not seen as a "monolithic" ability that "species either do or do not have" (Tomasello et al. 2003: 204), but rather "as configurations of features that constitute a family of related perspectivization capabilities of different degrees of complexity" (Verhagen 2008: 139).

Although the concepts of ToM and 'intersubjectivity' differ in several respects (see Zlatev et al. 2008), the gradual development of perspectival capacities is seen as significant in most of these approaches,³ as seen in the multi-stage-models proposed to account for the different degrees of perspectival tasks (cf. Perner 1991; Tomasello 1999: 179; Gärdenfors 2008; Zlatev 2008). All these models take for granted a crucial developmental step at the age of around four years.⁴ Whereas children under the age of four already manage shifts of perspective and hypothetical representations (Astington 1990: 157), it is at this time that children begin to pass the verbal test of 'false belief' that assesses whether they can attribute a false belief to another person and reason about the mental states of others (cf. e.g. Astington 1990; Leslie 2000; Rosenthal 2000; Gallagher & Hutto 2008). Second, this step is seen in close relationship to the development of more complex visual Level 2 perspectivization and linguistic perspectival tasks such as the usage of propositional attitude verbs, syntactic embedding, and grammatical means of epistemicity and evidentiality (cf. e.g. Harris 1996; Leslie 2000; Papafragou 2001; Astington & Baird 2005; de Villiers 2005; Moll & Meltzoff 2011a; b; c; Leiss 2012) as well as narrative comprehension (cf. e.g. Astington 1990; Feldman et al. 1990; Nelson 2003; Gallagher & Hutto 2008: 29; Goldie 2007).

In order to address the question of what triggers the increase of perspectival complexity, it seems thus promising to look at the common features between complex linguistic and cognitive perspectivization tasks. The basic idea in the following is thus that if we "break down both language and theory of mind into more basic components" (Lohmann, Tomasello & Meyer 2005: 263), this will offer insights with respect to the basic principles of perspectivization shared by language and cognition and their different degrees of perspectival complexity.

² Cf. the different (families of) theories such as 'theory theory', 'simulation theory', and 'interaction theory' (see Gallagher & Varga 2015 for an overview). In the following, ToM is understood as a term for different perspectival tasks of belief understanding without being specifically committed to one of these theories. While consequences for a theoretical account of ToM are not at issue in this paper, the multi-stage-model proposed in Section 5 would call for a pluralist approach as argued for by Gallagher & Varga (2015).

³ But see the two-system approach by Apperly & Butterfill (2009), who claim that the development of ToM is based on two cognitive systems that are conceptually radically different (cf. Carruthers 2016 for an excellent critical overview).

⁴ As a reaction on the study of Onishi & Baillargeon (2005), it has been controversially discussed whether false belief is a preverbal cognitive ability independent of language which is already present before the fourth year. While the question has not found a conclusive answer so far, the study shows the need for breaking down the degrees of perspectivization into more basic components.

3 The basic components of C(ognition)-perspectivization

In order to determine the basic components of C-perspectivization a view on developmental studies seems instructive where the gradual unfolding of perspectival abilities is described in multiple stages. According to these models, a child's viewpoint is at the primary level linked to the *here-and-now* of the real situation, so that “[t]he perceiver has no option of representing anything but current reality” (Perner 1991: 66f.). So while young children are able to share affective and perceptual experiences in terms of intersubjectivity and joint attention (cf. e.g. Gallagher & Hutto 2008: 21; Zlatev 2008: 224; Moll & Meltzoff 2011b: 394), such early forms of interpersonal interaction do not allow any detachment from the real context situation. A relevant step in the acquisition of perspectival capabilities is hence the mental decoupling from the real world, as seen in children's capabilities of differentiating between present and past as well as switching between actual and imaginary hypothetical situations. This could be seen as a form of *perspective-taking* since the decoupling of reality allows for alternate viewpoints: a banana can be seen either as a banana in the real world or a telephone in an imaginary setting as well as the child can choose to be a prince or a princess or someone else in a play. Children are thus able to select one possibility out of multiple possible situation models by choosing either one option or another (cf. Perner 1991: 50).⁵

While children at this age are thus capable of taking perspectives, they yet experience difficulties in more complex perspectival tasks such as the verbal false belief test.⁶ This standard test relies on an experimental set-up with two characters A and B and an object which is located at a specific place X. When A leaves the room, B moves the object from X to another place Y. When A comes back, the testee is asked to predict where A will look for the object and where the testee thinks that A believes the object is. In their answer, four-/five-year-olds usually consider the state of knowledge of the character who has left the room and point at the location where the object is meant to be according to A's knowledge. Younger children, however, tend to point at the location where the object really is according to their own knowledge. These different reactions are not trivial to explain since children younger than four to five years of age are able to leave their own perspective behind, and even to apprehend that others may see things differently (cf. Moll & Meltzoff 2011a; b; c).

According to Moll & Meltzoff (2011b; c), the intricacy in managing the false belief task does not lie in the “the ability to take another's perspective on an object” (Moll & Meltzoff 2011c: 299) but in the more challenging competence “to *confront* two (or more) perspectives on the selfsame object at the same time” (Moll & Meltzoff 2011b: 403; emphasis in original). Referring to Flavell's terminology, Moll & Meltzoff thus draw a distinction between Level 2a and 2b of perspective-taking, whereby only level 2b requires confronting perspectives (Moll & Meltzoff 2011c: 299). The false-belief test requires this more complex level 2b, since the testee has to handle two diverging systems of knowledge: she

⁵ Whether pretend plays are only based on dual representations or require metarepresentational capacities (and, as such, more than ‘perspective-taking’) is of controversial discussion. While they display an affinity to metarepresentational tasks in presupposing the awareness of pretending and, as such, a self-conscious reflection, it is a well-known fact that pretend plays are observed in children much earlier than the fourth year. Furthermore, it has been argued that pretend plays do not necessarily require second-order representations (Whiten & Byrne 1991). Jarrold et al. (1994) have shown that the controversy cannot be solved in a simple way since different types and levels of pretend play have to be distinguished, whereby some kinds of pretense are not metarepresentational. Against this background, I restrict the discussion in the following to the specific cases of imaginary context shifts as argued in Perner (1991).

⁶ The false-belief test as a standard setting for testing ToM capabilities has been criticized by e.g. Bloom & German (2000) and Fenici (2011) who argue that the test does not show whether the testee ‘has’ a ToM or not, but refers to a more specific mental perspectivization task. This is in line with the argument in the following that ToM comprises a rather complex perspectival structure.

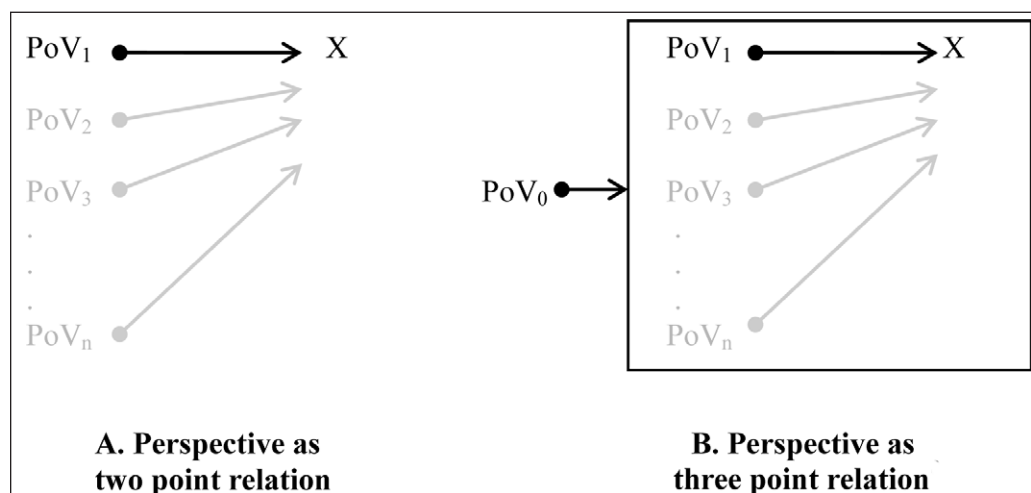


Figure 1: Concept of perspective as two-point vs. three-point relation.⁸

knows that A thinks that the object is in location X, while B knows that the object is in location Y – and she knows at the same time that only one possibility is correct, since she knows where the object really is. ToM tasks thus “require the ability to simultaneously represent conflicting information: the protagonist’s (or one’s) own belief about a past situation and the current true situation” (Plaut & Karmiloff-Smith 1993: 70). In order to give the right answer, she does not only have to suppress the own state of knowledge but needs to evaluate the two conflicting perspectives from her external stance.

With respect to the concept of perspectivization, this has important implications. First, a model of perspective as a two-point-relation becomes insufficient since it has to be complemented by a third reference point from which the contrasting viewpoints are evaluated, cf. Figure 1.⁷

Second, the three viewpoints are not equivalent but display a hierarchical difference, since the third reference point is superordinated to the other alternate ones. The hierarchy of the viewpoints requires the simultaneous activation of both levels of the viewpoint constellation and their relationship to each other. According to Perner (1991), perspectivization is thus closely linked to metarepresentation as “a representation of a representation as a representation” (Perner 1991: 35), since it represents the representational relation itself:⁹

[...] a child capable of mental metarepresentation who, for instance, represents that a picture is a representation needs to construct a mental model containing two substructures and their relationship. One structure has to represent the picture (as

⁷ Similar conclusions can be drawn with respect to the ‘Smarties test’ where children are confronted with an unexpected content. This test also requires to hold active one’s own present view (content = pencils) and the former state of belief (content = sweets) as well as the different knowledge systems of the testee and another person who does not know yet that the box does not contain sweets.

⁸ Note that the ‘three-point’ refers to a viewing situation between an observer, the observed entity and the external view on the relation in-between. As such, it is not the same as ‘joint attention’, i.e. a communication process of (at least) two people focusing on the same object of conversation which is commonly described in terms of ‘triangulation’ but does not require any higher-order intentions. In terms of perspective-taking, Moll & Metzloff (2011b: 288) refer to this stage as Level 0. In Davidson’s terms, two stages of ‘triangulation’ are distinguished – a primary stage which corresponds to early ‘joint attention’, and a second stage which presumes a second-order concept of ‘objectivity’. In the following, I will use the neutral term ‘three-point-relation’ in order to refer to the hierarchical viewpoint situation as described above.

⁹ Note that metarepresentation according to this definition is not the same as ‘a representation of a representation’ (see Perner 1991: 35; Scott 2011 for a distinction of different conceptions of metarepresentation within ToM accounts).

a physical entity) and the other what the picture depicts (its interpretation), and, very importantly, the model has to include links between these two structures representing how the picture relates to the depicted. (Perner 1991: 83)

For complex perspectivization tasks it is thus not sufficient to select one out of different equivalent viewpoints but to hold the whole viewpoint constellation in mind by evaluating contrasting alternatives from an external reference point with respect to their relationship to each other. As a result, the relationship between the different viewpoints itself becomes the object of the perspectivization process. The observations on the development of cognitive perspectivization tasks thus show that what such tasks require is “not perspective *taking* but the explicit acknowledgement that a given object may be seen in alternative ways” (Moll & Meltzoff 2011b: 408).

The ‘full’ concept of perspective thus comprises different degrees of perspectival complexity (cf. Table 1): The primary prerequisite is the givenness of more than one viewpoint that offers equivalent alternatives (i). While perspective-taking implies that one viewpoint is picked out by rejecting the other possible options in the sense of *either – or* (i.e. *either taking the banana as a banana or as a telephone*), more complex perspectival task display more complex constellations regarding the relationship between the viewpoints involved. This is already the case in choosing between real and hypothetical situations, where the viewpoint shift involves a decoupling process from the actual situation and can in this respect be paralleled to Flavell’s Level 1 of visual perspectivization; yet, the original viewpoint is not totally cancelled but suppressed as seen in the fact that children are much aware that a banana is not a telephone in the real world. That is to say that the primary origo remains active in the background as a reference point (ii). This is the prerequisite for a hierarchical constellation of perspectival embedding of viewpoints at different levels since it presupposes that one viewpoint is superior to the other (iii). Yet, the complexity of ToM tasks as the false belief test requires more than taking the perspective of another person. In the ToM task, one has to evaluate the diverging perspectives (of Sally, Ann and her own) with respect to their truth values, whereas the truth value in imaginary scenarios is quite clear. The processing effort of different perspectives thus requires the reflection about the constellation of viewpoints (in analogy to Level 2b of visual perspective taking in terms by Moll & Meltzoff 2011a; b). The qualitative difference between the viewpoints is linked to the emergence of a metarepresentational awareness with respect to the given potential of possible alternatives, i.e. an external third reference point arises from which the diverging viewpoints are evaluated and integrated in an overall picture, which corresponds to Level 2b of visual perspective taking (iv). This is also reflected by the results of recent studies on non-verbal tasks of false belief test (Onishi & Baillargeon 2005). As shown in Rubio-Fernández & Geurts 2013, children are able to pass the test (in the test design with a success rate of 80%) if the children are encouraged to continuously track the perspective of the protagonist, while any kind of disruption risks the failing at the test. This indicates that the test becomes more complicated if the mono-perspective is disturbed.

In sum, several degrees of complexities can thus be distinguished, cf. Figure 2.

These different degrees of perspectival complexity can also be observed in the emergence of pictorial perspectivization tasks. As Doherty & Wimmer 2005 have shown, the perception of ambiguous pictures such as the famous rabbit/duck representation by Wittgenstein is at first restricted to one perspective only: Children perceive the picture as *either* a duck *or* a rabbit, even if they are instructed about the ambiguity of the picture (Doherty & Wimmer 2005: 408f.). It is only later that they experience them like adults as oscillating between two possible interpretations. According to the authors, this is based on the same capabilities that are also necessary for more complex perspectival tasks, since it requires the understanding of the representational relationship between the figure and

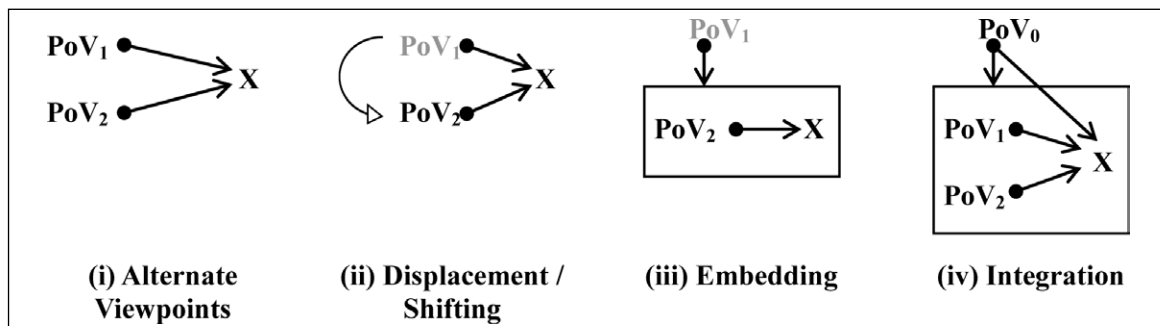


Figure 2: Degrees of perspectival complexity.

its two interpretations (Doherty & Wimmer 2005: 418). In a comparable way, drawings by young children represent individual objects from one particular perspective only. Entities like fishes for example, are commonly depicted in profile view, while lizards are drawn as seen from above. Such early pictures thus juxtapose different viewpoints that would be seen as contrasting from an overall integrating viewpoint. Yet, the different perspectives are not perceived as conflicting. It is in fact a rather late development that children integrate the different perspectives under one superordinated viewpoint like it is the case for linear perspective (cf. Köller 2004: 54).

It thus becomes clear that also with respect to pictorial representations perspectivization is more than perspective-taking in the sense of a selection of a certain viewpoint that restricts the aspects of the object in focus. My claim in the following will be that for L(anguage) perspectivization, the same mechanisms of viewpoint embedding and viewpoint integration as outlined in Figure 2 are necessary in order to account for viewpoint constellations in grammar and narrative discourse.

4 L-perspectivization

Against the background of the previous section, it will be shown that complex perspectival phenomena of L-perspectivization such as propositional attitude ascriptions, grammatical means of epistemicity and evidentiality, and narrative comprehension share similar viewpoint constellations and are based on the mechanism of *confronting* perspectives.

4.1 Viewpoint constellation of propositional attitude ascriptions

It is a commonplace that language as “an elaborated metarepresentational device” (de Saussure 2013: 68) is closely interrelated with cognitive capacities such as perspective-taking and Theory of Mind, and various studies have argued for a correlation between linguistic abilities of propositional attitude ascriptions and the performance on ToM tasks such as the false-belief test (cf. for overviews Astington & Baird 2005a; Jacques & Zelazo 2005: 144–146; Paynter & Peterson 2010). While it is controversial which capacities exactly can be paralleled with ToM capabilities, there is a close link between the structural components of ToM tasks and syntactic embedding under non-factive mental state verbs. This parallelism is already indicated in the definition by Dunbar (2006: 172):

Theory of mind is the ability to mind read or imagine how another individual sees the world. It is encapsulated in the statement: “I believe that you think the world is flat.”

The analogy seems quite obvious since complement clauses like the one in (1)

(1) I believe that you think the world is flat.

seem to reflect ToM as the ability to imagine another individual’s mental state insofar as the embedding structure of two mental state predicates puts two different subjects

of consciousness on stage (i.e. *I believe* vs. *you think*). This suggests that there is a viewpoint switch from the speaker-I to another person's stance, and, hence, perspective-taking. However, there are two aspects that go beyond a perspective switch. First, *taking* another person's stance does not mean to take another person's position and forget about one's own. Rather, it is 'taken' while maintaining the relation to one's own state of belief. The primary viewpoint and the second viewpoint are hence not equivalent options, but display a qualitative difference whereby one viewpoint is subordinated to the other. Second, both primary and secondary viewpoints are based on the complement structure of propositional attitude ascriptions: non-factive verbs of perception, cognition, or communication in the matrix clause indicate the perspective on the event, while the embedded clause contains the perspectivized entity (cf. Astington & Baird 2005b: 165; Verhagen 2005: 78). This is seen in the fact that utterances like (2) can be evaluated with respect to two different veridicality attributions but that only (2b) would be a natural reaction in discourse.

- (2) He believes that p [the world is flat]
 modus **dictum**
 a. "Yes, he does." (referring to the character's intensional state)
 b. "No, it is not." (referring to the truth value of the embedded proposition)

This complement structure is yet not a given. When children start producing sentences like *I think p* between the age of 2 and 3 years, mental state predicates are first used like epistemic stance markers (Diessel & Tomasello 2001). The subordination of the complement under the main clause is thus a later development. As Verhagen argues, it is only when this ability has developed that "it also becomes possible for a conceptualizer, in uttering *I think*, to construe his own thinking as an object of conceptualization for specific purposes, as in *I think he will arrive on time, but I am not sure/but John is skeptical [...]*" (Verhagen 2007: 71). This is a crucial step since it allows for the detachment of one's own viewpoint and to consider one's own perspective (i.e. the relation between the origo and the perspectivized entity) as an object of evaluation. As a result, the own viewpoint becomes the origo and (part of) the object of evaluation at the same time. This recursive structure calls for a three-point concept of perspectivization, since the evaluation of the own viewpoint presupposes the awareness that there are other possible perspectives available, in other words: that p [the world is flat] could be false. According to de Villiers et al. (2014: 229), it is thus the recognition that the complement is false and hence the contrast of truth-values between clauses that triggers the recursion of complementation structures.

Note that the observations so far hold for first order sentences like in (2) which refer to one subject of consciousness only. In (1), the structure is even more complex since it involves the embedding of another person's mental state and thus multiplies the relations between the different viewpoints. There is (i) a potential contrast between the two speaker beliefs (*I believe* vs. *you think*), and (ii) a potential contrast with respect to the embedded proposition that can be either true or false. While both contrasts are relevant with respect to the utterance, they do not display the same perspectival constellation. Contrast (i) between *The world is flat* vs. *The world is not flat* refers to two possible options that are structurally equal, but contradict each other in that only one of them can be true at the same time (see also Dancygier & Sweetser 2005; Dancygier 2012b for modelling the concept of 'alternativity' in mental space theory). The difference between *I believe* and *you believe*, on the other hand, relies on a contrast, but not a contradiction, as trivially, I can believe something other than you do – in the same way as it may be raining in Paris and not raining in New York. According to Perner et al. (2003), 'real' perspective problems are hence defined by the fact that the different viewpoints cannot be aligned by coordination.

- (3) a. The world is flat and the world is not flat. → contradicting
 b. I believe that the world is flat, and you believe that it is not. → contrasting

The concept of perspective-*taking* as defined in Section 3 presupposes that there are (at least) two alternative equal viewpoints to choose from. That is to say, there is a choice between *either* option A *or* option B, as given in (3a). The structure in (3b) integrates both the contradiction on the propositional content level as well as the contrast between the two states of mind which shows that embedding structures can integrate more than one viewpoint at the same time. This becomes clear by the fact that the second viewpoint can be filled out by an external ‘speaker’ which is not explicitly marked within the sentence, cf. the example in (4).

- (4) Little Red Riding Hood
 a. believes that **the wolf** is her grandmother.
 b. believes: ***The wolf** is my grandmother.
 → internal perspective (PoV = Little Red Riding Hood)
 → **external perspective** (PoV = speaker)

In (4), Little Red Riding Hood believes that the one lying in the bed is her grandmother, whereas the reader knows that this is not in line with the reality of the story world. The denotation of the one lying in the bed as “the wolf” thus cannot reflect the perspective of Little Red Riding Hood, but an external viewpoint which has to be reconstructed from the communicative context. This is seen in the fact that the denotation as “the wolf” is not possible in the direct speech construction in (4b) that reflects Little Red Riding Hood’s perspective. (4a) thus integrates two different perspectives: the viewpoint of Little Red Riding Hood whose belief system does not contain the fact that the one lying in her grandmother’s bed is the wolf, and an external point of view which contains the knowledge about the fact that the one lying in bed is actually the wolf. The latter is not explicitly marked in the linguistic structure but can be reconstructed from ‘outside’.

At this point, it is important to be precise about the notion of *internal* vs. *external*. As applied to example (4), the distinction refers to the structural difference between two embedded viewpoints. In this sense, internal₁ captures the fact that the viewpoint is a second, displaced viewpoint that is viewed from outside by another external₁ viewpoint that is situated at a different level. This distinction is independent of the fact whether the viewpoint is linked to e.g. proprioceptive, kinaesthetic, and cognitive processes ‘inside’ one’s mind, i.e. to private feelings and thoughts that can be accessed directly via introspection (e.g. ‘I have a headache’) (i.e. internal₂). As will be seen in Section 4.3, both distinctions are relevant with respect to the perspectival structure but can lead to different evaluations of the viewpoint constellation.

In sum, it is thus trivially true that (1) is perspective-*taking* in the sense that a speaker uttering it chooses a certain perspective. But that is not the interesting point. Rather, the look on propositional attitude ascriptions shows that linguistic perspectivization is more than perspective-taking for more than one reason. First, as pointed out in the literature (Dancygier & Sweetser 2012; Dancygier & Vandelanotte 2016), even rather straightforward examples display not only one, but different viewpoints. Second, as also shown by Dancygier & Vandelanotte (2016), these viewpoints display hierarchical contrasts. With respect to the constellation of viewpoints, it is thus necessary to differentiate between equal alternatives that allow for perspective-taking in terms of an *either or* distinction (i.e. stage (ii) in Figure 2) and vertical relations of embedding (iii). Third, the integration of viewpoints requires a third external reference point from which the potentially contrasting viewpoints can be evaluated (iv). This calls for a three-point concept of perspective. As will be shown in the following section, the same is true for perspectivization in grammar.

4.2 Viewpoint constellation of epistemic modality

In a wide sense of the term, grammar is genuinely perspectival in that every grammatical paradigm offers a choice of alternatives. In a more narrow sense, grammar is linked to perspectivization insofar as it locates the speaker's deictic origo that determines the 'view' of the verbal event situation (cf. the approaches by Leiss 1992 and Diewald 2010 in tradition of Bühler and Jakobson). In cognitive linguistics, this is captured in the premise that "[i]nherent in every usage event is a presupposed viewing arrangement, pertaining to the relationship between the conceptualizers and the situation being viewed" (Langacker 2001: 16). It follows from this that grammatical markers do not indicate a certain viewpoint, but rather a viewpoint constellation (see Verhagen 2016; Dancygier & Sweetser 2012; Sweetser 2012; Dancygier & Vandelanotte 2016 for examples). This can be illustrated by a look at the category of tense. At first glance, tense markers like the simple past are "shifters" (Jakobson 1957) since they relocate the reference point from a present utterance time to the past. Hence, they could be seen as instances of perspective-taking. Moreover, such a displacement of the origo in the sense of "Deixis am Phantasma" (Bühler 1934 [1999]) relies on the persistence of the primary origo in the background. The temporal perspectivization thus refers to the relation between primary and secondary reference points and, hence, a multiperspectival constellation between viewpoints of different qualities. As such, it requires a three-point description, as reflected in the ternary system of Reichenbach 1947 that models tense as relations between the time of event and the time of speech as perspectivized by a third reference point (cf. in detail Zeman 2015; see also Evans 2005 with respect to the multiperspectivity of complex tenses).

Nevertheless, tense markers are prototypical shifters insofar as they shift the focus on the displaced viewpoint whereas the primary origo is only maintained in the background. This is different for markers of evidential and epistemic modality that display a more complex perspectival structure. As laid out in the introductory section, there seems to be a strong link between the development of ToM capabilities and the comprehension of epistemic and evidential modality that is a rather late development in language acquisition (Perner 1991: 150; Papafragou 1997: 16; Papafragou 2001; Leiss 2012). This becomes understandable by the fact that epistemic and evidential meanings share with propositional attitude ascriptions the feature that they are modifiers of propositions, and as such, overt markers of speakers' attitudes. Both evidentiality and epistemicity thus display a similar perspectival structure by embedding veridical statements that may be true or false. This is seen in the fact that only (5b) but not (5a) can be translated into the structure of an epistemic modal verb.

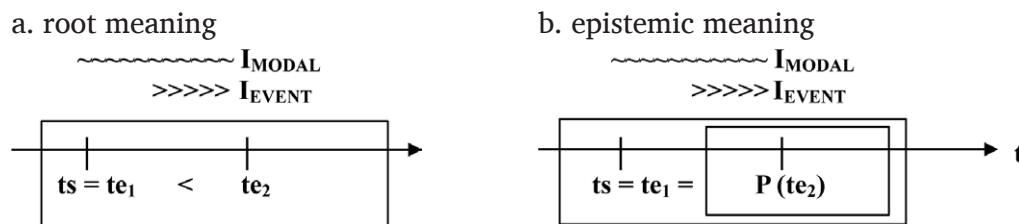
- (5) a. I think of cake and wine.
 *That must be cake and wine!
 b. I think that the one lying in her grandmother's bed is the wolf.
 The one lying in her grandmother's bed must be the wolf!
 I think that p [the creature is the wolf]

In contrast to (5a), (5b) leads to the potential for viewpoint contrasts between the level of the sentence subject and the level of the speaker, and as such, the potential for an internal₁ and an external₁ view in the structural sense (i.e. inside vs. outside the proposition). This is also a crucial difference in comparison to the perspectival structure of root modals, cf. (6).

- (6) a. Little Red Riding Hood must walk through the woods in order to see her grandma.
→ ‘She is obliged to do so.’
- b. Little Red Riding Hood must be walking through the woods right now.
→ ‘I (i.e. the speaker) guess that she is doing so right now’
(but reality could teach us otherwise).

Root modals like in (6a) are characterized by their biphasicness, since they refer to two different time intervals, the temporal interval of the modal in the real world (i.e. the time for which the obligation holds) and the time interval of the event denoted by the infinitive complement in a possible world. As such, they display the potential for a focus shift from the present viewpoint to the future event, as seen in the fact that modals constitute a grammaticalization source for future tenses (cf. Zeman 2013). The main difference with respect to the epistemic example (6b) is yet that it introduces the speaker’s viewpoint, and, hence, an additional potential contrast between two belief systems (external₂ vs. internal₂): while the obligation in (6a) holds for the subject, i.e. Little Red Riding Hood, the modal meaning in (6b) scopes over the whole proposition *p* [Little Red Riding Hood is walking through the woods] and thus requires an external viewpoint of evaluation.

- (7) Biphasic structure of (present tense) modal verbs:



(I = intervall; ts = speech time; te = event time; P = proposition)

The perspectival complexity of (7b) is thus based both on the structural embedding of viewpoints (i.e. inside vs. outside the proposition), the distinction between two belief contents, and a contrast of veridicality on the propositional level, i.e. the fact that the proposition could be false. So once again, the construction integrates external and internal viewpoints at the same time. As such, they display a pattern of viewpoint integration (i.e. pattern (iv) in Table 1; Level 2b in the terminology by Moll’s & Meltzoff’s 2011). The difference in degree of complexity between (7a) and (7b) is also reflected in language acquisition. As laid out in Section 4.1, complementation structures like *think that p* are a later development than constructions like *think of p* that can be analysed in terms of perspective-taking. In analogy, verbs of volition are used prior to verbs of belief, and the usage of root modals precedes the comprehension of epistemic and evidential modals (cf. de Villiers 2005; Perner et al. 2005). In sum, it is not the potential viewpoint switch – and thus: not perspective-taking – that is crucial with respect to the difference between both constructions, but the viewpoint constellation that is based on the contrast between two different hierarchical levels (internal₁ vs. external₁), two different belief systems (internal₂ vs. external₂), and the contrast with respect to the veridicality attribution on the propositional level (true vs. false).

4.3 Viewpoint constellation in narrative discourse

The degrees of perspectival complexity are also identifiable for perspectivization in narrative contexts. Narratives are perspectival structures par excellence since they integrate viewpoints of different characters and narrators and thus offer a set of possible alternate perspectives that allow for viewpoint switches within the text. However, also narrative perspectivization is more than perspective-taking for at least two reasons. First, the different viewpoints are not equivalent alternatives since the viewpoints of characters and narrators can be embedded in each other (cf. Dancygier 2012a; Dancygier & Vandelanotte 2016). This is in particular relevant with respect to the structural difference between narrator and character level since narration is characterized by the fact that the narrator level scopes over the event level of the protagonists (cf. also Zeman 2016; to appear).¹⁰ Second, these different viewpoints require in addition a global perspective from which the different perspectives are organized (see also Dancygier 2012; this volume; Dancygier & Vandelanotte 2016):

[...] viewpoints are organized hierarchically and in terms of a network, with local viewpoint choices achieving overall coherence in what one might call a top-level or ‘Discourse Viewpoint’ space, from which lower-level viewpoint choices are overseen. (Dancygier & Vandelanotte 2016: 14)

As a result, different layers of perspectivization arise. This can be best illustrated by a look at Free Indirect Discourse (FID), a perspectival pattern that is restricted to narrative discourse mode only and thus allows for some insights with respect to the structure of narration in general, cf. (8).

(8) She was glad. Tomorrow was finally the day she was to see her granny again!

As (8) shows, FID blends together two different viewpoints (cf. Pascal 1977; see also Fludernik 1993: 316ff. for an overview; Vandelanotte 2009: 246–251 for discussion of ‘dual voice accounts’). While deictic elements such as *tomorrow* and the emotive choice of kinship terms like *granny* match the character’s perspective, personal pronouns and tenses are linked to the level of the narrator and thus allow for the reconstruction of the narrator’s viewpoint (cf. in detail Schlenker 2004; Sharvit 2008; Maier 2012; Eckardt 2014). The relationship between these two perspectives has been accounted for in terms of double context and is preferably analysed as (partial) context shift. In this respect, Eckardt (2014) has proposed to evaluate sentences in FID relative to a pair of contexts: “an external context *C* that is shared by the narrator and the reader, and an internal context *c* corresponding to a situation in which a protagonist of the story talks or thinks, as part of the story” (Eckardt 2014: 60).

As laid out before, the ‘internal’ vs. ‘external’ can refer to two different aspects, namely (i) the hierarchical difference between the communicative levels of discourse, whereby the narrator has naturally an ‘outside’ view on the character on the discourse level (whether this character is referentially himself or not) and (ii) the question of whether the contents on the propositional level are ‘thoughts’ that can be accessed directly (‘internally’). Due to its functional ‘outside’ position, the perspective of the narrator allows for simultaneous knowledge about his commitment towards the proposition and the course of the story which can include the mental contents of the protagonists. The narrator knows what the protagonist knows, while the reverse would lead to metaleptic structures. In a structural sense then, FID integrates both internal₁ and external₁ viewpoints at the same time.

¹⁰ This hierarchical difference is independent from the fact whether the narrator is fleshed out on the discourse level as a figural instance or remains covert.

On the other hand, however, FID is also characterized by the fact that on the discourse level, only one perspective is foregrounded, i.e. the character's point of view, while the narrator's voice remains 'invisible' and can only be traced back within the grammar, i.e. the use of tenses and pronouns.¹¹ As a result, the reported propositional content in (8) is perceived as dependent on the character's viewpoint – which could be reliable or not. So while FID structurally integrates two different perspectives (i.e. internal₁ vs. external₁), it is at the same time committed to the perspective of just one subject of consciousness that is restricted to mental contents of speech and thought (i.e. internal₂) (cf. Vandelanotte 2009: 246–251 for the argument that FID could be considered both as “bivocal” and “univocal”). As a result, grammatical constructions that are bound to the narrator's perspective are incompatible with FID. This can be exemplified by the German 'future of fate' (FoF) construction that displays the complementary perspectival effect of FID:

- (9) Am nächsten Tag **sollte** Rotkäppchen seine Großmutter wieder **sehen**.
 'The next day, Little Red Riding Hood was to see (literally: 'should see') her grandma again.'

While FID foregrounds the perspective of the character, the modal verb construction in (9) restricts the perspective to the narrator's viewpoint. He knows what will happen next, whereas the character on the story level is unaware of the events to come. Such a view from outside is incompatible with the character perspective which seems to be the cause why the FoF-reading is blocked in FID (cf. Eckardt 2012: 185). In a sentence like *She was glad. Tomorrow she should see her again!* the event to come is focalized from the character's perspective, linked to a less degree of certainty than the proleptic prediction by the narrator. The hierarchy of perspectives is seen in the fact that the character is subordinate to the force of the narrator's point of view whose controlling authority is grounded on its functional role within the narrative discourse. The authority of the narrator makes the event realization highly certain and rules out any intervention of the character on the story level. Hence, both constructions – FID and FoF – act on the double context of narration and structurally integrate a potential contrast of viewpoints, but the viewpoint constellation is reverse in that one perspective is suppressed in favor of the other on the discourse level.

So is this perspective-taking? Yes and no. Apart from the trivial fact that it is perspective-taking on the side of the speaker who selects a certain perspective in choosing either FID or FoF (or any other option), one could argue that it is perspective-taking in the sense that the perspective *either* of the narrator *or* the character is 'taken'. On the other hand, it is more than just perspective-taking, since both (8) and (9) refer to more than one subject of consciousness (i.e. internal₂ PoVs that are linked to mental agents) and rely on a hierarchical contrast between external₁ and internal₁ PoVs (in the structural sense of the term).

Furthermore, both constructions are based on a contrast between the narrator's and the character's viewpoint. In FoF, the divergent states of knowledge trigger a narrative tension since the narrator (and the reader) knows more than the character about the things to come. In FID, the report from the character's perspective could turn out as unreliable or create ironic effects in contrast to the narrator's viewpoint. In both cases, the potential contrasts thus call for a third evaluating viewpoint on a global level from which the perspectives are monitored. This is in line with Klein & von Stutterheim (2002) who distinguish between local perspectives within the text and “a globally established perspective” (Klein & von Stutterheim 2002: 83) that maintains the *quaestio* of the story. With respect

¹¹ The narrator's invisibility has even led to the fact that FID has been a major argument in challenging the view that every narrative necessarily presumes a narrator (cf. e.g. Banfield 1982; Fludernik 1993: 443; Patron 2009).

to narrative discourse structure in particular, Dancygier (2012a), this volume, has argued that narratives are structured in different narrative spaces that require “a high-level story-viewpoint (SV) space, which structures the text-wide viewpoint of the story” (Dancygier 2012a: 129f.). That is to say, a global viewpoint is required that monitors the different perspectives in order to establish a coherent view on the story.

The establishment of such a global viewpoint is also a fundamental developmental step, as it is well known that stories of young children commonly lack a coherent macro-organization (cf. von Stutterheim & Carroll 2007: Fn. 3, with reference to Berman & Slobin 1994; see also Büttner 2016 with respect to disorders of macro-structural planning from a neuro-linguistic perspective). This does not only affect the production, but also the reception of stories. According to Goldie (2007), the capability of taking an “external perspective”, i.e. thinking about another person in terms of “‘he thinks that p’ and ‘he thinks that if p then q’” (Goldie 2007: 70), constitutes the prerequisite for the appreciation and evaluation of diverging perspectives in narratives (cf. similarly Zunshine 2006 with respect to the role of ToM-task in narrative appreciation). Thereby, the comprehension of ‘dramatic irony’ displays more than a shift of perspective since the reader has to take into account (at least) two diverging viewpoints simultaneously in order to appreciate the story: the reader has to understand that the one in bed that Little Red Riding Hood is seeing is actually the wolf while knowing at the same time that Little Red Riding Hood thinks that it is her grandmother. Dramatic irony thus displays the same viewpoint constellation as laid out in Section 3 for false belief understanding: The beliefs of two mental subjects as alternate viewpoints that have to be evaluated from a third reference point as contrasting ones with respect to their propositional content and evaluated with respect to their relation to each other.

5 Modeling perspectival complexity in language and cognition

The comparison between propositional attitude ascriptions, epistemic modality, and narrative discourse structure has thus shown that the different linguistic phenomena display the same perspectival principles as complex cognitive perspectivization tasks. For all instances, it would thus be too simplifying to describe the perspectivization process as *taking* a perspective in the sense of selecting one viewpoint out of several ones available that restricts the aspects of the object seen. Rather, it is the constellation of potentially contrasting viewpoints that is relevant with respect to linguistic perspectival phenomena, whether we focus on grammar, complement structure or narrative discourse. Based on these observations, we can thus derive the following scale of perspectival complexity as reflected in both language and cognition (cf. also Table 1 below).

(i) Alternate viewpoints

For both cognitive as well as linguistic perspectivization, the crucial prerequisite is the givenness of alternate, i.e. divergent viewpoints. This constitutes the basis for perspective-taking in the sense of selecting one option out of equivalent possible alternatives in terms of *either – or*.

(ii) Viewpoint shift

As soon as there is more than one viewpoint available, the possibility of viewpoint switch arises. This mechanism is reflected in hypothetical representations as well as instances of direct discourse (Perner et al. 2002: 1466) and the difference between present and past that require a decoupling from the ‘real world model’ (Perner 1991) and the *hic et nunc* origo in Bühler’s (1934 [1999]) term of ‘Deixis am Phantasma’.

Such forms of ‘shifting’ (Jakobson 1957) and ‘displacement’ (Bühler 1999[1934]) constitute already more complex tasks since the principle of ‘Deixis am Phantasma’

(Bühler 1934 [1999]) implies (i) that the original viewpoint is not necessarily cancelled but maintained, and (ii) that the real primary and the secondary displaced origo are not equal viewpoints on the same level but dependently related to each other. As such, they provide the potential for viewpoint embedding.

(iii) Viewpoint embedding

The qualitative difference between primary and secondary viewpoints can lead to a hierarchical difference when one viewpoint is in the scope of another, as for example seen above in the distinction between character and narrator level in narrative discourse. This leads to a structural difference between internal₁ and external₁ viewpoints, in other words, “between a metarepresentation *about* a ‘belief world’ and a simulated assertion made ‘from within’ a belief world which we assume” (Recanati 2000: 61). ‘Shifts’ between hierarchical viewpoints are thus more complex than perspective-taking since they are based on a recursive structure where the external perspective includes the internal one by definition.

(iv) Viewpoint integration

Whereas viewpoint embeddings are more complex viewpoint constellations than viewpoint shifting, FID and FoF constructions, false-belief tasks, and narrative comprehension display an even more complex perspectival structure, since they are based on the maintenance of (at least) two diverging viewpoints and require the capability to process the diverging perspectives of internal_{1/2} and external_{1/2} contexts at the same time. This integration of (potentially) contrasting viewpoints requires a third external reference point from which the constellation is evaluated. It is thus again the process of *confronting perspectives* that lays the foundation for perspectival complexity.¹²



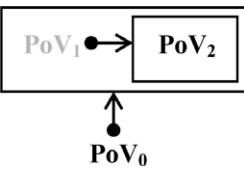
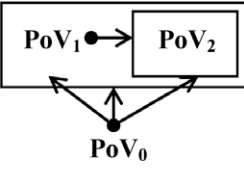
	C-perspectivization	L-perspectivization
	Alternate (divergent) viewpoints (either or)	Alternate (divergent) viewpoints (either or)
	Viewpoint shift Detachment of the ‘real world model’ (Perner 1991): past vs. present, real vs. hypothetical situations	Viewpoint shift ‘Deixis am Phantasma’: grammatical shifting (tense, mood)
	Embedding Internal vs. external perspective: second-order beliefs	Embedding Internal vs. external perspective: propositional attitude ascriptions, grammatical shifting (epistemicity, evidentiality), character vs. narrator distinction
	Integration Simultaneous activation of internal and external viewpoints (and, at the same time): false belief, metarepresentation	Integration Simultaneous activation of internal and external viewpoints (and, at the same time): FoF construction, Free Indirect Discourse, ‘dramatic irony’

Table 1: Degrees of perspectival complexity in C- and L-perspectivization.

¹² These steps are also compatible with the model by Dancygier (2012a) who distinguishes in terms of mental space theory between projection, embedding, and blending. Thereby, the difference between embedding and blending is seen in the fact that embedding is a result of projection where the hierarchical organization is maintained (cf. Dancygier 2012a: 190). By difference, blended spaces level any syntactic dependence.

6 Conclusion: Perspectivization as a three-point structure

At the beginning of the chapter, we started with the rather trivial and uncontroversial fact that perspectivization can be seen as a selection of a viewpoint that restricts the aspects of the object in focus. In this sense, perspectivization seems at first to be a matter of *either or*: a house is seen *either* from outside *or* inside, with respect to *either* its front *or* its back. This is true for perceptual viewing situations in the real world. As it comes to mental, pictorial, and linguistic conceptualization, however, perspectivization is more than just taking one perspective. In this respect, it has been shown that the different degrees of perspectival complexity are determined by (i) the qualitative difference between alternate viewpoints that are linked to a structure of subordination (e.g. primary real vs. second displaced viewpoints) and (ii) the establishment of a third reference point of evaluation that allows to maintain more than one perspective on different hierarchical levels at the same time. The main point of the paper has thus been that, if we metaphorically map the notion of perspective on mental, pictorial, and linguistic representation systems, we have to take into account not only the two-point-relation between an observing subject and an observed entity, but also the meta-relation that integrates the two-point-relation(s) under a third reference point as an evaluating instance.

As a consequence, phenomena of L- and C-conceptualization such as false belief understanding, propositional attitude ascriptions, epistemic modality, and narrative discourse have been characterized by the fact that they maintain external and internal perspectives at the same time and thus require the capability of confronting perspectives. In this respect, they could be compared to cubistic paintings that juxtapose different contrasting viewpoints of one object on the canvas. What makes L-perspectivization even more complex than cubistic pictures is yet the fact that, as seen above, it is necessary to distinguish between different local and global levels in order to account for the perspectival viewpoint constellations. L-perspectivization thus appears to be more comparable to the works of M.C. Escher that integrate local deviant perspectives which cannot be true at the same time but yet constitute coherent pictures as seen from a global level. So, in sum, perspectivization seems to capture nothing less than the relational architecture of language. Seen from this viewpoint, perspectivization is indeed fundamental.

Abbreviations

C-perspectivization = perspectivization in cognition, FID = free indirect discourse, FoF = future of fate construction, I = interval, L-perspectivization = perspectivization in language, PoV = point of view, ToM = theory of mind, ts = speech time, te = event time, P = proposition

Acknowledgement

The author is thankful to the editors and four anonymous reviewers for their thoughtful suggestions on a previous version of this paper.

Competing Interests

The author has no competing interests to declare.

References

- Apperly, Ian & Stephen Butterfill. 2009. Do humans have two systems to track beliefs and belief-like states? *Psychological Review* 116(4). 953–970. DOI: <https://doi.org/10.1037/a0016923>
- Astington, Janet Wilde. 1990. Narrative and the child's theory of mind. In Bruce K. Britton & Anthony D. Pellegrini (eds.), *Narrative thought and narrative language*, 151–171. Hillsdale, NJ: Erlbaum.

- Astington, Janet Wilde & Jodie A. Baird. 2005a. Introduction: Why language matters. In Janet Wilde Astington & Jodie A. Baird (eds.), *Why language matters for theory of mind*, 3–25. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195159912.003.0001>
- Astington, Janet Wilde & Jodie A. Baird. 2005b. Representational development and false-belief understanding. In Janet Wilde Astington & Jodie A. Baird (eds.), *Why language matters for theory of mind*, 163–185. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195159912.001.0001>
- Astington, Janet Wilde & Jodie A. Baird (eds.). 2005. *Why language matters for theory of mind*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195159912.001.0001>
- Banfield, Anne. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan Paul.
- Berman, Ruth & Dan I. Slobin. 1994. *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale: Erlbaum.
- Bloom, Paul & Tim P. German. 2000. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* 77. B25–B31. DOI: [https://doi.org/10.1016/S0010-0277\(00\)00096-2](https://doi.org/10.1016/S0010-0277(00)00096-2)
- Bühler, Karl. 1999 [1934]. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Reprint Jena: Fischer.
- Büttner, Julia. 2016. Neurolinguistic view into narrative processing. In Natalia Igl & Sonja Zeman (eds.), *Perspectives on narrativity and narrative perspectivization* (Linguistic Approaches to Literature 21), 63–88. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/lal.21.04but>
- Carruthers, Peter. 2016. Two systems for mindreading? *Review of Philosophy and Psychology* 7(1). 141–162. DOI: <https://doi.org/10.1007/s13164-015-0259-y>
- Dancygier, Barbara. 2012a. *The language of stories*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511794414>
- Dancygier, Barbara. 2012b. Negation, stance, and subjectivity. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 69–96. Cambridge: Cambridge University Press.
- Dancygier, Barbara. 2012c. Conclusion: multiple viewpoints, multiple spaces. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 219–231. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/cbo9781139084727.016>
- Dancygier, Barbara & Eve Sweetser. 2005. *Mental spaces in grammar: Conditional constructions*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486760>
- Dancygier, Barbara & Lieven Vandelandotte. 2016. Discourse viewpoint as network. In Barbara Dancygier, Wei-lun Lu & Arie Verhagen (eds.), *Viewpoint and the fabric of meaning form and use of viewpoint tools across languages and modalities*, 13–40. Berlin/New York: de Gruyter. DOI: <https://doi.org/10.1515/9783110365467-003>
- de Saussure, Louis. 2013. Perspectival interpretations of tenses. In Louis de Saussure & Kasia Jaszczolt (eds.), *Time: Language, cognition & reality*, 46–69. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199589876.003.0004>
- de Villiers, Jill. 2005. Can language acquisition give children a point of view? In Janet Wilde Astington & Jodie A. Baird (eds.), *Why language matters for theory of mind*, 186–219. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195159912.003.0010>
- de Villiers, Jill, Kathryn Hobbs & Bart Hollebrandse. 2014. Recursive complements and propositional attitudes. In Thomas Roeper & Margaret Speas (eds.), *Recursion: Complex-*

- ity in cognition* (Studies in Theoretical Psycholinguistics 43), 221–242. Cham: Springer. DOI: https://doi.org/10.1007/978-3-319-05086-7_10
- Diessel, Holger & Michael Tomasello. 2001. The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12. 97–141. DOI: <https://doi.org/10.1515/cogl.12.2.97>
- Diewald, Gabriele. 2010. On some problem areas in grammaticalization studies. In Katerina Stathi, Elke Gehweiler & Ekkehard König (eds.), *Grammaticalization: Current views and issues*, 17–50. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/slcs.119.04die>
- Doherty, Martin J. & Marina C. Wimmer. 2005. Children's understanding of ambiguous figures: Which cognitive developments are necessary to experience reversal? *Cognitive Development* 20. 407–421. DOI: <https://doi.org/10.1080/17405629.2010.515885>
- Dunbar, Robin I. M. 2006. Brains, cognition and the evolution of culture. In Stephen C. Levinson & Pierre Jaisson (eds), *Evolution and culture*, 169–180. Cambridge MA: MIT Press.
- Eckardt, Regine. 2012. Who is talking? – Narrators and speakers in free indirect discourse. Ms. URL: <https://www.uni-goettingen.de/de/323434.html> (25 Dec 2016).
- Eckardt, Regine. 2014. *The semantics of free indirect discourse: How texts allow us to mind-read and eavesdrop*. Leiden: Brill. DOI: <https://doi.org/10.1163/9789004266735>
- Evans, Nick. 2005. View with a view: Towards a typology of multiple perspective constructions. *Annual Meeting of the Berkeley Linguistics Society* 31/1. 93–120. DOI: <https://doi.org/10.3765/bls.v31i1.3429>
- Farrant, Brad M., Janet Fletcher & Murray T. Maybery. 2006. Specific language impairment, theory of mind, and visual perspective taking: Evidence for simulation theory and the developmental role of language. *Child development* 77(6). 1842–1853. DOI: <https://doi.org/10.1111/j.1467-8624.2006.00977.x>
- Feldman, Carol Fleisher, Jerome Bruner, Bobby Renderer & Sally Spitzer. 1990. Narrative comprehension. In Bruce K. Britton & Anthony D. Pellegrini (eds.), *Narrative thought and narrative language*, 1–78. Hillsdale, NJ: Erlbaum.
- Fenici, Marco. 2011. What does the false belief test test? *Phenomenology and Mind* 1. 197–207.
- Flavell, John H. 1992. Perspectives on perspective taking. In Harry Beilin & Peter B. Pufall (eds.), *The Jean Piaget symposium series. Vol. 14. Piaget's theory: Prospects and possibilities*, 107–139. Hillsdale, NJ: Erlbaum.
- Flavell, John H., Barbara Everett, Karen Croft & Eleanor R. Flavell. 1981. Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology* 17(1). 99–103. DOI: <https://doi.org/10.1037/0012-1649.17.1.99>
- Fludernik, Monika. 1993. *The fictions of language and the languages of fiction: The linguistic representation of speech and consciousness*. London: Routledge.
- Gallagher, Shaun & Daniel D. Hutto. 2008. Understanding others through primary interaction and narrative practice. In Jordan Zlatev, Timothy P. Racine, Chris Sinha & Esa Itkonen (eds.), *The shared mind: Perspectives on intersubjectivity* (Converging evidence in language and communication research 12), 17–38. DOI: <https://doi.org/10.1075/celcr.12.04gal>
- Gallagher, Shaun & Somogy Varga. 2015. Social cognition and psychopathology: A critical overview. *World Psychiatry* 14. 5–14. DOI: <https://doi.org/10.1002/wps.20173>
- Gärdenfors, Peter. 2008. Evolutionary and developmental aspects of intersubjectivity. In Hans Liljenström & Peter Århem (eds.), *Consciousness transitions: Phylogenetic, ontogenetic and physiological aspects*, 281–305. Amsterdam: Elsevier. DOI: <https://doi.org/10.1016/B978-044452977-0/50013-9>

- Goldie, Peter. 2007. Dramatic irony, narrative, and the external perspective. In Daniel D. Hutto (ed.), *Narrative and understanding persons*, 69–84. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/S1358246107000045>
- Graumann, Carl F. 1989. Perspective setting and taking in verbal interaction. In Rainer Dietrich & Carl F. Graumann (eds.), *Language processing in social context*, 93–122. Amsterdam: North Holland. DOI: <https://doi.org/10.1016/b978-0-444-87144-2.50007-0>
- Harris, Paul L. 1996. Desires, beliefs, and language. In Peter Carruthers & Peter K. Smith (eds.), *Theories of theories of mind*, 200–220. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511597985.014>
- Jacques, Sophie & Philip David Zelazo. 2005. Language and the development of cognitive flexibility: Implications for theory of mind. In Janet Wilde Astington & Jodie A. Baird (eds.), *Why language matters for theory of mind*, 144–162. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195159912.003.0008>
- Jakobson, Roman. 1971 [1957]. Shifters, Verbal categories and the Russian Verb. In Roman Jakobson (ed.), *Selected writings II*, 130–147. The Hague: Mouton.
- Jarrold, Chris, Peter Carruthers, Peter K. Smith & Jill Boucher. 1994. Pretend play: Is it metarepresentational? *Mind and Language* 9. 445–468. URL: <http://faculty.philosophy.umd.edu/pcarruthers/Pretend-play.htm> (30-sep-2016).
- Klein, Wolfgang & Christiane von Stutterheim. 2002. Quaestio and L-perspectivation. In Carl F. Graumann & Werner Kallmeyer (eds.), *Perspective and perspectivation in discourse* (Human cognitive processing 9), 59–88. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/hcp.9.06stu>
- Köller, Wilhelm. 2004. *Perspektivität und Sprache. Zur Struktur von Objektivierungsformen in Bildern, im Denken und in der Sprache*. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110919547>
- Langacker, Ronald. 2001. Cognitive linguistics, language pedagogy, and the English present tense. In Martin Pütz, Susanne Niemeier & René Dirven (eds.), *Applied cognitive linguistics: Theory and language acquisition*, 3–40. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110866247>
- Leiss, Elisabeth. 1992. *Die Verbalkategorien des Deutschen. Ein Beitrag zur Theorie der sprachlichen Kategorisierung* (Studia Linguistica Germanica 31). Berlin: de Gruyter.
- Leiss, Elisabeth. 2012. Epistemicity, evidentiality, and Theory of Mind (ToM). In Werner Abraham & Elisabeth Leiss (eds.), *Modality and Theory of Mind elements across languages* (Trends in Linguistics. Studies and Monographs 243), 39–65. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110271072>
- Leslie, Alan M. 2000. How to acquire a representational Theory of Mind. In Dan Sperber (ed.), *Metarepresentations: A multidisciplinary perspective*, 197–224. Oxford: Oxford University Press.
- Lohmann, Heidemarie, Michael Tomasello & Sonja Meyer. 2005. Linguistic communication and social understanding. In Janet Wilde Astington & Jodie A. Baird (eds.), *Why language matters for theory of mind*, 245–265. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195159912.003.0012>
- Maier, Emar. 2012. Quotation and unquotation in free indirect discourse. Ms. Groningen. URL: <https://sites.google.com/site/emarmaier/publications> (30-sep-2016).
- Michelon, Pascale & Jeffrey M. Zacks. 2006. Two kinds of visual perspective taking. *Perception & Psychophysics* 68(2). 327–337. DOI: <https://doi.org/10.3758/BF03193680>
- Moll, Henrike & Andrew N. Meltzoff. 2011a. How does it look? Level 2 perspective-taking at 36 months of age. *Child development* 82(2). 661–673. DOI: <https://doi.org/10.1111/j.1467-8624.2010.01571.x>

- Moll, Henrike & Andrew N. Meltzoff. 2011b. Joint attention as the fundamental basis of understanding perspectives. In Axel Seemann (ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience*, 393–413. Cambridge, MA: MIT Press.
- Moll, Henrike & Andrew N. Meltzoff. 2011c. Perspective-taking and its foundation in joint attention. In Johannes Roessler, Hemdat Lerman & Naomi Eilan (eds.), *Perception, causation, and objectivity*, 286–304. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199692040.003.0016>
- Nelson, Katherine. 2003. Narrative and the emergence of a consciousness of self. In Gary D. Fireman, Ted E. McVay & Owen J. Flanagan (eds.), *Narrative and consciousness: Literature, psychology, and the brain*, 17–36. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195140057.003.0002>
- Onishi, Kristine H. & Renée Baillargeon. 2005. Do 15-month-old infants understand false beliefs? *Science* 308(5719). 255–258. DOI: <https://doi.org/10.1126/science.1107621>
- Papafragou, Anna. 1997. Modality in language development: a reconsideration of the evidence. *UCL Working Papers in Linguistics* 9. URL: <http://papafragou.psych.udel.edu/papers/uclwp97.pdf> (30-sep-2016).
- Papafragou, Anna. 2001. Linking early linguistic and conceptual capacities: The role of Theory of Mind. In Alan Cienki, Barbara J. Luka & Michael B. Smith (eds.), *Conceptual and discourse factors in linguistic structure*, 169–184. Chicago: The University of Chicago Press.
- Pascal, Roy. 1977. *The dual voice. Free indirect speech and its functioning in the nineteenth-century European novel*. Manchester: Manchester University Press.
- Patron, Sylvie. 2009. *Le Narrateur. Introduction à la théorie narrative*. Paris: Armand Colin.
- Paynter, Jessica & Candida Peterson. 2010. Language and ToM development in autism versus Asperger syndrome: Contrasting influences of syntactic versus lexical/semantic maturity. *Research in Autism Spectrum Disorders* 4. 377–385. DOI: <https://doi.org/10.1016/j.rasd.2009.10.005>
- Perner, Josef. 1991. *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Perner, Josef, Johannes L. Brandl & Alan Garnham. 2003. What is a perspective problem? Developmental issues in belief ascription and dual identity. *Facta Philosophica* 5. 355–378.
- Perner, Josef, Sandra Stummer, Manuel Sprung & Martin Doherty. 2002. Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development* 17(3–4). 1451–1472. DOI: [https://doi.org/10.1016/S0885-2014\(02\)00127-2](https://doi.org/10.1016/S0885-2014(02)00127-2)
- Plaut, David C. & Annette Karmiloff-Smith. 1993. Representational development and theory-of-mind computations. *Behavioral and Brain Sciences* 16. 70–71. DOI: <https://doi.org/10.1017/S0140525X0002906X>
- Recanati, François. 2000. *Oratio obliqua, oratio recta: An essay on metarepresentation*. Cambridge, MA: MIT Press.
- Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press.
- Rosenthal, David M. 2000. Consciousness and metacognition. In Dan Sperber (ed.), *Metarepresentation in evolutionary perspective*, 265–299. Oxford: Oxford University Press.
- Rubio-Fernández, Paula & Bart Geurts. 2013. How to pass the false-belief task before your 4th birthday. *Psychology Science* 24(1). 27–33. DOI: <https://doi.org/10.1177/0956797612447819>
- Schlenker, Philippe. 2004. Context of thought and context of utterance (A note on free indirect discourse and the historical present). *Mind and Language* 19(3). 279–304. DOI: <https://doi.org/10.1111/j.1468-0017.2004.00259.x>

- Scott, Sam. 2011. Metarepresentation in philosophy and psychology. In Johanna D. Moore & Keith Stenning (eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, 910–916.
- Sharvit, Yael. 2008. The puzzle of Free Indirect Discourse. *Linguistics and Philosophy* 31(3). 353–395. DOI: <https://doi.org/10.1007/s10988-008-9039-9>
- Surtees, Andrew, Ian Apperly & Dana Samson. 2013. Similarities and differences in visual and spatial perspective-taking processes. *Cognition* 129(2). 426–438. DOI: <https://doi.org/10.1016/j.cognition.2013.06.008>
- Sweetser, Eve. 2012. Introduction: Viewpoint and perspective in language and gesture, from the ground up. In Barbara Dancygier & Eve Sweetser (eds.), *Viewpoint in language: A multimodal perspective*, 1–22. Cambridge: Cambridge University Press.
- Tomasello, Michael. 1999. *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael, Josep Call & Brian Hare. 2003. Chimpanzees versus humans: It's not that simple. *Trends in Cognitive Sciences* 7. 239–240. DOI: [https://doi.org/10.1016/S1364-6613\(03\)00107-4](https://doi.org/10.1016/S1364-6613(03)00107-4)
- Vandelanotte, Lieven. 2009. *Speech and thought representation in English: A cognitive-functional approach* (Topics in English Linguistics 65). Berlin: de Gruyter.
- Verhagen, Arie. 2005. *Constructions of intersubjectivity: Discourse, syntax, and cognition*. Oxford: Oxford University Press.
- Verhagen, Arie. 2007. Construal and perspectivization. In Dirk Geeraerts & Herbert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 48–81. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199738632.013.0003>
- Verhagen, Arie. 2008. 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* (Converging Evidence in Language and Communication Research 12), 307–331. DOI: <https://doi.org/10.1075/celcr.12.17ver>
- Verhagen, Arie. 2016. Introduction: On tools for weaving meaning out of viewpoint threads. In Barbara Dancygier, Wei-lun Lu & Arie Verhagen (eds.), *Viewpoint and the fabric of meaning form and use of viewpoint tools across languages and modalities*, 1–10. Berlin/New York: de Gruyter. DOI: <https://doi.org/10.1515/9783110365467-002>
- von Stutterheim, Christiane & Mary Carroll. 2007. Durch die Grammatik fokussiert. *Zeitschrift für Literaturwissenschaft und Linguistik* 145. 35–60. DOI: <https://doi.org/10.1007/BF03377187>
- Whiten, Andrew & Richard W. Byrne. 1991. The emergence of metarepresentation in human ontogeny and primate phylogeny. In Andrew Whiten (ed.), *Natural Theories of Mind*, 267–282. Oxford: Blackwell.
- Zeman, Sonja. 2013. Zur Diachronie der Modalverben: *Sollen* zwischen Temporalität, Modalität und Evidentialität. In Werner Abraham & Elisabeth Leiss (eds.), *Funktionen von Modalität* (Linguistik – Impulse und Tendenzen 55), 335–366. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110333091.335>
- Zeman, Sonja. 2015. The elementary particles of distance in space, time, grammar, and discourse. In Barbara Sonnenhauser & Anastasia Meermann (eds.), *Distance in language: Grounding a metaphor*, 7–36. Newcastle Upon Tyne: Cambridge Scholars Publishing.
- Zeman, Sonja. 2016. Perspectivization as a link between narrative micro- and macro-structure. In Sonja Zeman & Natalia Igl (eds.), *Perspectives on narrativity and narrative perspectivization* (Linguistic Approaches to Literature 21), 17–42. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/lal.21.02zem>

- Zeman, Sonja. To appear. What is a narration – and why does it matter? In Markus Steinbach & Annika Hübl (eds.), *Linguistic foundations of narration in spoken and sign language* (Linguistics Today/Linguistik Aktuell). Amsterdam: Benjamins.
- Zlatev, Jordan. 2008. The co-evolution of intersubjectivity and bodily mimesis. In Jordan Zlatev, Timothy P. Racine, Chris Sinha & Esa Itkonen (eds.), *The shared mind: Perspectives on intersubjectivity* (Converging Evidence in Language and Communication Research 12), 215–244. DOI: <https://doi.org/10.1075/celcr.12.13zla>
- Zlatev, Jordan, Timothy P. Racine, Chris Sinha & Esa Itkonen (eds.). 2008. *The shared mind: Perspectives on intersubjectivity* (Converging Evidence in Language and Communication Research 12). Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/celcr.12>
- Zlatev, Jordan, Timothy P. Racine, Chris Sinha & Esa Itkonen (eds.). 2008. Intersubjectivity: What makes us human. In Jordan Zlatev, Timothy P. Racine, Chris Sinha & Esa Itkonen (eds.), *The shared mind. Perspectives on intersubjectivity* (Converging Evidence in Language and Communication Research 12), 1–14. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/celcr.12.02zla>
- Zunshine, Lisa. 2006. *Why we read fiction: Theory of mind and the novel*. Columbus, OH: Ohio State University Press.

How to cite this article: Zeman, Sonja. 2017. Confronting perspectives: Modeling perspectival complexity in language and cognition. *Glossa: a journal of general linguistics* 2(1): 6.1–22, DOI: <https://doi.org/10.5334/gjgl.213>

Published: 09 February 2017

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS The Open Access logo, which is a stylized 'a' inside a circle.