

Stef Spronck\* and Tatiana Nikitina\*

# M and R as elements of a syntactic unit: Where would the relation between M and R come from, if not from syntax?

<https://doi.org/10.1515/lingty-2019-0014>

First of all we would like to thank all the commentators for their valuable contributions. They provide additional evidence for our analysis, but they also challenge aspects of it in interesting ways.

Unfortunately, space prevents us from addressing most of the points raised here. As Hodge and Cormier demonstrate, discussions about reported speech quickly touch upon fundamental questions about the role of convention in linguistic description and the contribution of multimodality. Dancygier highlights the importance of finding specific theoretical answers for data we encounter in the realm of reported speech, which cannot be easily treated with familiar notions such as ‘family resemblances’. And McGregor shows the exciting prospect that our observations can be taken even further than we had anticipated. We are also grateful for the avenues of debate Maier has opened up between a typological approach and the rich literature in formal semantics on quotation, and we thank Quer for doing the same with the more recent but equally rich literature on sign languages.

Much to our delight, there seems to be a rather general (though not complete) agreement among the commentators that reported speech is a relatively coherent phenomenon with distinct properties. However, there is variation in their interpretation of the level of analysis at which this phenomenon should be treated. In this response we would like to briefly address one question that emerged in several commentaries, particularly that of Rumsey, Reesink, Goddard and Wierzbicka, and Güldemann (who even chooses it as the title of his commentary): What, exactly, is *syntactic* about reported speech?

After outlining our answer to this question in Section 1, we sketch how this approach fits within a broader research programme (Section 2), and wrap up with some concluding remarks (Section 3).

---

**\*Corresponding authors: Stef Spronck [stef sprɔŋk]**, University of Helsinki, P.O. Box 24 (Unioninkatu 40), 00014 Helsingin Yliopisto, Helsinki, Finland, E-mail: stef.spronck@helsinki.fi  
**Tatiana Nikitina [tatʲˈjanə nʲiˈkʲitʲinə]**, CNRS, LLACAN, 7, rue Guy Môquet - BP 8, 94801 Villejuif, France, E-mail: tavnik@gmail.com

# 1 Syntactic typology and the standing Bakhtinian challenge of reported speech

Despite the great variety of views voiced by the commentators, most criticism focuses on one central aspect of our account: the notion of syntax. While almost none of the commentators disagree with our claim that reported speech shows some degree of coherence in the set of conventional features it is associated with, few of the commentators follow us in attributing it a common syntactic nature. Pointing to the variation of formal patterns observed in the expression of reported speech, Dancygier asks “why [...] should [we] insist on referring to this broad spectrum of instances as a ‘syntactic domain’, given that the phenomena within the scope of the definition are not quite syntactically cohesive, while being cohesive semantically.” In a similar vein, Rumsey observes that “it seems paradoxical that although their criteria for delimiting the domain of reported speech are all ultimately semantic, their primary aim is to establish that reported speech “constitutes a dedicated *syntactic* domain”.”

These comments draw attention to an aspect of our analysis that, we agree, was left too implicit in our target paper: our proposal can only work if syntax is conceived of as a domain involving multiple levels of abstraction. Our syntactic category of reported speech represents the most abstract of these levels, encompassing more specific syntactic levels, such as clausal constructions in individual languages. When we suggest, for example, that elements take on a different meaning in the context of R, this may not only apply to features or particles, but equally to, e.g. a complement clause. In many languages reported speech involves structures that undeniably resemble subordinate clauses or, as in Rumsey’s examples, transitive constructions, but they are rather idiosyncratic instantiations of these constructions (as we observed in Section 2 of our paper). The reason for this is, we suggest, that they are instantiated as part of, or recruited by, an encompassing syntactic function, which we have characterised as a relation between an abstract element M and an abstract element R. Crucially, these elements are defined by their mutual relationship: an M is an M because it forms part of an abstract unit with an R, and *vice versa*. This status of reported speech, as an abstract relationship that does not derive from its lexical components, is what qualifies this function as syntactic.

The relative consensus among the commentators that reported speech can be meaningfully characterised as a broad level phenomenon, however, does not lead to a shared view that this phenomenon can therefore be characterised as syntactic. The commentators provide an array of possible alternative characterisations such as ‘Viewpoint network’ (Dancygier), functional domain

(Goddard & Wierzbicka), enactment (Hodge & Cormier) and recursive discourse (Reesink). We actually do not fundamentally object to any of these characterisations at some level of analysis. However, we would suggest that once we agree that, cross-linguistically, reported speech involves a more or less stable relationship between abstract elements that are expressed through a complex morphosyntactic structure, the conclusion that it therefore (at least also) involves syntax becomes rather trivial.

This brings us to a point Güldemann raises: we have not explicitly contrasted reported speech with other syntactic construction types. In the context of syntactic argumentation this is a valid question, but we would like to counter it with two comments: First, we would like to point to the work by McGregor (1994, 1997). McGregor demonstrates extensively that the syntactic relation involved in reported speech cannot be reduced to either coordination/parataxis or subordination/hypotaxis, two of the most prominent syntactic relations at the sentential level. Both the present account and Güldemann (2008) explicitly build on this analysis. Our arguments about the idiosyncrasy of M and R further add to the evidence that these units are not 'regular' main and sub-clauses, respectively.

Second, the commentators offer two specific construction types that in their analysis would be more profitably characterised as syntactic, instead of our 'reported speech'. We hope that by explicating how these structures relate to our view of reported speech as a syntactic category, we can also clarify the relevance of exploring syntactic contrasts of the type Güldemann suggests.

The first challenge is formed by the set of structures Rumsey brings up in his observations about transitivity and the 'individuation' of R. Specifically, Rumsey states that in several languages R is marked as an object clause, even receiving accusative marking in the Australian language Martuthunira (Dench 1995). The second proposed candidate is the opposition between direct and indirect speech, which Reesink and Goddard & Wierzbicka suggest as the more appropriate syntactic level at which to discuss reported speech in typology. Reesink observes that in Usan, subtle referential properties distinguish between what could be labelled a direct and an indirect speech construction. Reesink and Rumsey both accept our definition as a reasonable unifying characterisation of reported speech expressions in their data, but they would only recognise the transitive constructions and direct/indirect speech constructions as syntactic units, and not reported speech itself.

We do not find this position consistent. First of all, it assumes that syntactic analyses at the language-specific level of, e.g. transitive constructions in the Papuan and Australian examples, are incompatible with the cross-linguistically applicable definition we presented. But would it be sufficient to say that, e.g. in Martuthunira there are subject-object clause constructions that somehow take

on a reported speech meaning? Or would it be more consistent to say that, given the semantic similarity of these constructions with reported speech elsewhere, reported speech in Martuthunira recruits subject-object constructions for its expression? We would suggest the latter option: the category of reported speech is a prerequisite for recognising the specific subject-object clause constructions in Martuthunira as an instantiation of reported speech.

Object complement clauses are, for us, syntactic constructions that are recruited in individual languages in order to express the abstract syntactic category of reported speech. These are language-specific and stand in paradigmatic opposition with other constructions within the given language.<sup>1</sup> The abstract relation of reported speech, however, applies across languages (which does not mean that it is universal in the sense that all languages should have it) and is not defined on the basis of oppositions with language-specific syntactic categories. If reported speech is a syntactic category at this ‘high’, abstract level, this helps to further explain the rather extreme range of semantic diversity associated with what appear to be reported speech constructions (as discussed in section 2.7 of our target paper): these constructions can be recruited for the expression of reported speech, but may also be recruited for the expression of other categories or features.

We propose that a similar analysis applies to Reesink’s suggestion that direct/indirect speech should be taken as the basis for a typological syntactic discussion of the phenomenon of reported speech. Evans (2013) certainly shows that it may be possible to construct direct/indirect speech as a canonical category/comparative concept, but doing so serves a goal that is different from ours: it aims to characterise a subtype of reported speech in semantic terms. Again, saying that direct and indirect speech constructions are relevant syntactic classes in individual languages is not incompatible with recognising reported speech as a cross-linguistic class, but it does not address the more fundamental point that we raise: Across the entire plethora of constructions associated with reported speech, they seem to have more in common with each other, than they have with the specific construction types they instantiate. In language after language we see that sentential constructions of reported speech *resemble*, for instance, complement clauses, but on closer inspection they display properties that are not common to complement clauses at all (as we have illustrated in Section 2 of our target article). The same applies to morphological expressions of

---

<sup>1</sup> This interpretation also opens up the possibility that a construction that contrasts with direct speech may actually not be an instantiation of reported speech itself. Note, for example, the interesting question Maier raises whether all instances of indirect speech are necessarily also instances of reported speech.

reported speech: yes, reportativity resembles a subclass of evidentiality in many languages, but actually it is also a very atypical subclass (Hengeveld & Hattner 2015; also see Maier's commentary). Our claim is simply that given these observations, it is not sufficient to state that these constructions are randomly strange, but that they share a fundamental property: they are part of a syntactic class we label reported speech. Focusing on the direct/indirect speech opposition simply sidesteps this issue.

Another question that came up in several commentaries concerns the analytical status of our definition: Is our syntactic category of reported speech a comparative concept in the sense of Haspelmath (2010)? Interestingly, the commentators take diverging views on this. Güldemann states that our approach does not amount to establishing a comparative concept, Reesink suggests that direct and indirect speech, rather than the more inclusive notion of reported speech, should be taken as the basis for a comparative concept, whereas Hodge and Cormier interpret our definition as a comparative concept itself.

We believe that Hodge and Cormier's reading is possible and that our definition could be interpreted as a comparative concept in the sense of Haspelmath (2010). In other words, it could be seen as a notional concept that is exclusively designed for the purposes of typological comparison and for diagnosing structures in individual languages. Our own interpretation would, however, be different. We agree with McGregor that reported speech is actually *more* than a comparative concept. It represents a distinctive constellation of meanings that can be expressed through a specific type of linguistic structures, and we believe that these properties uniquely identify reported speech as a category. The structural means available in a language to signal, e.g. demonstratedness, are highly language-specific, but they are not random: from the expressive means Quer identifies in sign languages,<sup>2</sup> to the deictic properties Reesink finds in Usan, demonstratedness can be related to structural (e.g. pronominal) and ad-hoc means of expression.

A consequence of this interpretation is that features that traditionally lie outside the realm of syntax, such as prosody, voice quality and dramatic gesture, which even commentators who highly value such forms of expression,

---

2 As we have stated, we agree with Quer that the term 'reported speech' is not entirely felicitous, to which he adds the further argument that it seems to leave out non-spoken languages. None of the alternatives proposed are without objections either, however. For example, reported 'discourse', which Quer suggests as a modality-neutral alternative, is colloquially primarily associated with written language/text, which removes associations with interactionality without introducing a connection with sign language. We see no objection to extending the notion of 'speech' to include 'sign language', and despite its downsides, the name 'reported speech' places our approach in a tradition going back to Vološinov (1973).

like Hodge & Cormier and Quer, call ‘pragmatic’ or types of communicative action can be compared with, e.g. clausal expressions of R. Rather than arguing for a syntax-centric view of language, this advocates an approach in which cross-modal expressions can be related to core properties of linguistic convention: demonstratedness can form part of syntax and is an intrinsic component of reported speech.

## 2 The Bakhtinian/Vološinovian research programme on reported speech

We are not the first ones to draw attention to the potential of reported speech for expanding our understanding of syntax. This is a crucial aspect of the research programme put forward by Vološinov (1973).

Several commentators explicitly refer to Bakhtin/Vološinov (1973) in support of the idea that reported speech is central to culture and human being-in-the-world. We agree with this view. But the central message of the discussion of reported speech in Vološinov (1973) is another one. Yes, for Bakhtin/Vološinov (1973) reported speech reflects the dialogic nature of human existence. But within Bakhtinian philosophy *everything* does. Every aspect of language is shaped by the speaker in a way that anticipates the addressee’s response, and that responds to anything that has come before in the dialogue. Shouting ‘No!’, using definite rather than indefinite reference, nervously swallowing the end of a sentence in response to a stern look: none of these uses of language are more or less ‘dialogic’ than reported speech.

What makes reported speech particularly interesting is that it displays its dialogic properties, in Vološinov’s famous phrase, ‘in the stabilized constructional patterns of the language itself’ (Vološinov 1973: 116). In other words, reported speech is dialogue in syntactic form, allowing us to study the dialogic properties of language at the level of conventional structure (see also Spronck Submitted). As a consequence, characterising the syntax of reported speech is an entry point for understanding the relation between social interaction and morphosyntax within this programme, rather than an ultimate goal (despite Hodge and Cormier’s characterisation of the work, Vološinov (1973) presents itself as a sociological approach to *grammar*, rather than communicative action).

We consider the approach sketched in our target paper a typological application of the Bakhtinian/Vološinovian account of reported speech. As such, we hope that it may prompt a more extensive examination of the view that reported speech constitutes an extraordinary linguistic structure, and that its linguistic

analysis can lend insight into how dialogic properties shape language and can be identified in language. Despite the abundant references to Bakhtin/Vološinov in the literature on reported speech, this fundamental idea has rarely been pursued in a consistent way (see Spronck Submitted for a more detailed discussion of this point).

### 3 Final observations and conclusion

The commentators have indicated several interesting pathways along which the discussion initiated here can progress, of which we would like to list a few.

Rumsey raises the topic of acquisition, and the gradual development in children of understanding reported speech as a coherent class. We do not see a contradiction between our claim that the reported speech consists of a pair of relatively discrete M and R elements and his observation that the fact that children acquire performative speech before more ‘complex’ types of RS suggests the value of a scalar approach between seeing R as a more or less individuate unit. This observation would suggest to us that until the language learner has construed a unit in which M and R are perceived as elements with a distinct role, the broader syntactic category of reported speech has not yet been acquired. Considering the diversity of the syntactic structures that are recruited for the expression of reported speech, the acquisition of reported speech as a cohesive grammatical domain is likely to be slow and gradual. The observation that young children often have difficulty distinguishing language-specific constructions associated with reported speech supports this view (Köder & Maier 2016).

McGregor adopts and extends our notion of defenestration far beyond the way in which we use the term. We agree with McGregor that we are surrounded by real and imagined speech events. A posture can be ‘telling’, a work of art ‘speaks to us’, a funny hat yells: “look at me”. Such examples, as Pascual (2007, 2014) and Pascual and Sandler (2016) demonstrate, are not incidental examples of literary English, they are everywhere. Thinking about the world in terms of how other people think about the world is what constitutes culture, and by implication, conversation provides an often used model for conceptualising interactions with objects or individual speech events (for example, Pascual (2007) convincingly argues that even rhetorical questions could be thought of as a way of recruiting a dialogic structure to express a ‘monologic’ meaning).

Although we are thrilled by McGregor’s exploration of defenestration beyond the more restrictive role we had imagined for it, we do feel that it is slightly counterproductive to extend the notion of ‘defenestrated clause’ to every

instance of Pascual's (2007, 2014) 'fictive interaction', including the ScotRail example. This is one of the main reasons why we believe evidentiality should count as a coded, rather than an implied meaning of reported speech. In order to exclude examples such as the ScotRail example from the class of reported speech constructions proper, Spronck (2016) argues that evidentiality constitutes a defining criterion, which can even be shown to identify, e.g. 'fictive' direct speech as a separate structural class in individual languages.

We are grateful to McGregor for showing several clear directions in which future discussions can develop. Maier hints at several further stimulating questions that such a dialogue may involve:

'Do cases of standard indirect discourse or hearsay evidentiality really always involve 'demonstratedness'? And in what sense does a free indirect speech or thought representation in a fictional narrative mark 'evidentiality' or 'source of information'?"

In conclusion, we would like to end with a general comment about contemporary debates about syntax, and a practical point for typology coming out of our proposal.

We would like to state that many objections voiced by commentators against syntax are involved in a 'shadow conversation' (Irvine 1996). Authors working within cognitive-functional and multimodal approaches to language understandably respond to a long history of neglect of language use, signed languages, gesture and (other) visual communicative signals and the (perceived) arrogance of syntax-centric accounts in much of the twentieth century. Our defence of the role of syntax should not be seen as a continuation of this discourse.

The poem in the epitaph accompanying Reesink's commentary illustrates this well: syntax is construed as a mechanical, computational system with no feeling, whereas reported speech is an emotion-laden, social phenomenon. Hodge and Cormier refer to syntax as 'symbolic', while language is embodied action. But why would syntax have to be this way?

Most objections against applying the notion of syntax to reported speech are a version of the argument put forward by D'Arcy (2015): traditional definitions of syntax cannot capture the richness and variety of reported speech and therefore syntax does not matter (or does not matter much). But this conclusion only follows if we accept that the traditional interpretation of syntax is *right*. We have attempted to argue that reported speech shows such a degree of idiosyncrasy and complexity at a cross-linguistic level, that the most consistent position is to analyse it as a syntactic category in itself, a category that can be defined as a relation between two abstract units M and R. Once we accept this, this category



needs to involve meanings not traditionally associated exclusively with syntax, such as ‘demonstratedness’ and indexicality, in the broader sense we have outlined. From this starting position we can further debate the suitability of these specific meaning components, and how generally they apply across languages. Through these properties, we suggest that reported speech offers an opportunity to break away from this model entirely and to explore syntax beyond the boundaries previously imposed.

Finally, our more practical point: if our suggestion is correct that reported speech constructions constitute a specific grammatical environment with its own constructional effects and processes of interpretation, we should treat examples of elements occurring *within* reported speech with suspicion, or at least with special care. Impressionistically, examples of reported speech are frequently used in descriptive grammars to illustrate phenomena such as verbal categories that happen to occur in R, modal particles, or other types of (inter)subjective grammatical markers. This is to be expected, since, particularly in the absence of actual dialogue data, reported speech may be the most prominent or even the only environment in which highly interactional elements show up in a corpus. If our assumption is that, e.g. direct speech constructions simply reflect ‘verbatim’ dialogue in an unfiltered way, this is unproblematic. However, if our suggestion is correct that elements, specifically those occurring in R, take on a special meaning due to the context of an R structure, observations about such elements should not automatically be generalised to contexts outside reported speech. Doing so may be particularly tempting in the case of defenestrated clauses, where no explicit marking of M is present.

Behind many of the arguments exchanged here lie different positions about the nature of language, the status and relevance of linguistic conventions, the appropriate level at which grammatical analysis should operate, and how to interpret the role of language structure in communicative action. But especially because this debate touches upon such fundamental questions we sincerely hope that this is not the end of the conversation. We would like to thank the commentators again for taking this first step with us.

## References

- D’Arcy, Alexandra. 2015. Quotation and advances in understanding syntactic systems. *Annual Review of Linguistics* 1(1). 43–61. doi:10.1146/annurev-linguist-030514-125220.
- Dench, Alan. 1995. *Martuthunira: A language of the Pilbara region of Western Australia*. Canberra: Pacific Linguistics.

- Evans, Nicholas. 2013. Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 66–98. Oxford etc.: Oxford University Press.
- Güldemann, Tom. 2008. *Quotative Indexes in African Languages: A synchronic and Diachronic Survey*. Berlin: Mouton de Gruyter.
- Haspelmath, Martin. 2010. [Comparative concepts and descriptive categories in crosslinguistic studies](#). *Language* 86(3). 663–687. doi:10.1353/lan.2010.0021.
- Hengeveld, Kees & Marize Mattos Dall’Aglío Hattner. 2015. Four types of evidentiality in the native languages of Brazil. *Linguistics* 53(3). 479–524.
- Irvine, Judith T. 1996. Shadow conversations: The indeterminacy of participant roles. In Michael Silverstein & Greg Urban (eds.), *Natural histories of discourse*, 131–159. Chicago/London: The University of Chicago Press.
- Köder, Franziska & Emar Maier. 2016. Children mix direct and indirect speech: Evidence from pronoun comprehension. *Journal of Child Language* 43(4). 843–866.
- McGregor, William B. 1994. The grammar of reported speech and thought in Gooniyandi. *Australian Journal of Linguistics* 14(1). 63–92.
- McGregor, William B. 1997. *Semiotic grammar*. Oxford: Clarendon Press.
- Pascual, Esther. 2007. Fictive interaction blends in everyday life and courtroom settings. In Todd Oakley & Anders Hougaard (eds.), *Mental spaces in discourse and interaction*, 79–107. Amsterdam/Philadelphia: John Benjamins.
- Pascual, Esther. 2014. *Fictive interaction: The conversation frame in thought, language, and discourse*. Amsterdam/Philadelphia: John Benjamins.
- Pascual, Esther & Sergeiy Sandler (eds.). 2016. *The conversation frame: Forms and functions of fictive interaction*. Amsterdam/Philadelphia: John Benjamins.
- Spronck, Stef. 2016. Evidential fictive interaction (in Ungarinyin and Russian). In Esther Pascual & Sergeiy Sandler (eds.), *The conversation frame: Forms and functions of fictive interaction*, 255–275. Amsterdam/Philadelphia: John Benjamins.
- Spronck, Stef. Submitted. Speaking for Bakhtin: two interpretations of reported speech (A response to Goddard & Wierzbicka 2018). *Russian Journal of Linguistics*.
- Vološinov, Valentin N. (ed.). 1973. *Marxism and the philosophy of language*. Ladislav Matejka & I.R. Titunik. New York/London: Seminar Press.