Stef Spronck* and Tatiana Nikitina*

Reported speech forms a dedicated syntactic domain

https://doi.org/10.1515/lingty-2019-0005
Received January 08, 2018; revised August 01, 2018

Abstract: In many languages, expressions of the type 'x said: "*p*"', 'x said that *p*' or 'allegedly, *p*' share properties with common syntactic types such as constructions with subordination, paratactic constructions, and constructions with sentence-level adverbs. On closer examination, however, they often turn out to be atypical members of these syntactic classes. In this paper we argue that a more coherent picture emerges if we analyse these expressions as a dedicated syntactic domain in itself, which we refer to as 'reported speech'. Based on typological observations we argue for the idiosyncrasy of reported speech as a syntactic class. The article concludes with a proposal for a cross-linguistic characterisation that aims at capturing this broadly conceived domain of reported speech with a single semantic definition.

Keywords: reported speech, quotation, reportativity, syntax, optionality, demonstration

1 Reported speech

1.1 Introduction

Expressions of reported speech, e.g. involving structures such as (1) and (2), tend to have several idiosyncratic features that distinguish them markedly from other construction types.

(1) John said: "Look, there is marmalade here!"

(2) John says that there was a typhoon yesterday.

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

For example, reported speech constructions often involve different kinds of 'deictic shift'. While in English constructions such as (1) the present tense 'is' and the demonstrative 'here' in the reported sentence typically indicate the perspective of the reported speaker (i.e. a deictic shift away from the current speaker), languages have been shown to make fine-grained distinctions in the elements that can undergo deictic shift (De Roeck 1994; Evans 2013), and across languages the patterning of pronouns in reported speech differs fundamentally from their behaviour outside reported speech contexts (Nikitina 2012a, Nikitina 2012b). Recent studies in newly documented languages have found discrepancies between the interpretation and marking of categories such as tense, modality and evidentiality within and outside reported speech (Gawne 2013; Munro et al. 2012; Spronck 2015a). And even though reported speech constructions may often resemble coordinated or appositional constructions, as in (1), or subordinate clauses, as in (2), they display several characteristics not normally associated with these sentence types, such as 'subordinated' clauses that are salient in terms of information structure, the length of the elements involved, which can be either greater or smaller than a clause, and atypical prosody (Spronck 2017; McGregor 1994).

Given these observations, in the present paper we would like to defend the following claim:

(3) Reported speech constitutes a dedicated syntactic domain, i.e. crosslinguistically it involves a number of specific/characteristic phenomena that cannot be derived from the involvement of other syntactic structures in reported speech, such as subordination.

The claim in (3) is in line with syntactic analyses proposed both in functional (Frajzyngier 1991; McGregor 1994, McGregor 1997; Spronck 2017; Vandelanotte 2008) and formal (Etxepare 2008; Speas 2004) paradigms that single out reported speech constructions as involving a dedicated syntactic relation that differs from other sentential structures. However, given the great variety of structures involved in the expression of reported speech, this claim is neither obvious, nor has it, to our knowledge, been given sufficient attention in typology so far.[1] Our goal here will be to identify and classify phenomena occurring in the context of reported speech, and to propose benchmarks for establishing reported speech as a cross-linguistic category.

1 A notable exception to this general characterisation is Güldemann (2008), a regional typological study discussing languages with a highly grammaticalised type of elements like 'he said' as in (1) under the notion of 'quotative indexes'. Our discussion here is indebted to his account, and also, especially, to the theoretical account of Gooniyandi reported speech by McGregor (1994) on which both the present paper and Güldemann (2008) build.

In order to further specify and substantiate the general claim in (3), we begin in Section 1.2 by providing a working definition of reported speech. While this working definition will soon prove inadequate, it should give a first indication of the phenomena we will include in our analysis. In Section 2 we then introduce eight fundamental observations about reported speech that we believe strongly support our claim about the syntactic exceptionality of reported speech, and we also illustrate the type of phenomena that require explanation. In Section 3 we propose a typological definition of reported speech. We suggest that this definition may serve as a cross-linguistic description of the wide range of reported speech structures and can account for the features observed in Section 2. Finally, Section 4 offers a brief conclusion.

1.2 A working definition of reported speech

The phenomenon we refer to here as reported speech has received various labels in the literature. Tannen (2007) introduces the notion of 'constructed discourse', highlighting the fact that when talking about speech events, the current speaker necessarily abstracts away from the 'original' speech situation and is thereby always ultimately responsible for constructing the form of the 'report' (cf. Plank 1986: 285). With a similar motivation, other authors have used the term 'represented speech' (McGregor 1997; Vandelanotte 2004; Verstraete 2011), underlining that what the speaker in utterances such as (1) and (2) intends is to give a 'representation' of the speech used, not to accurately 'report' it (also cf. Banfield 1982). A third notion that is found in typological descriptions of the phenomenon is 'quotation' (cf. Evans 2013); although in its common interpretation this term may be relatively neutral, it has received by far the most varied interpretations in the literature. For example, Clark and Gerrig (1990) use the notion exclusively to refer to 'direct speech' (i.e. constructions like the one in (1)), and the label has a long history in formal semantics and the philosophy of language, where it can apply to the act of referring to language more generally, and to other expressions that can be written between quotation marks (or 'mentioned' rather than 'used'), such as mentioning a book title (Saka 1998). As this description suggests, the quotation literature mainly centres on written (European) languages (but see Ginzburg & Cooper 2014).

Given the ambiguity inherent in the term 'quotation' in the literature we have avoided it here, and have opted for 'reported speech'. We will adopt this more traditional notion, which goes back to the first translation of Vološinov (1930), a classic treatment of reported speech, with its often cited definition as 'speech within speech, utterance within utterance, and at the same time also *speech about speech, utterance about utterance*' (italics in the original) (Vološinov 1973: 115).

The term, alongside its close equivalent 'reported discourse,'[2] has also been used as a label in the four main typologically oriented volumes on the topic (Coulmas 1986a; Güldemann & von Roncador 2002; Janssen & van der Wurff 1996; Lucy 1993), as a well as in countless descriptive and typological studies (Aaron 1992; Besnier 1993; Bogomolova & Ganenkov 2010; Bugaeva 2008; Creissels 2010; Ershova 2012; Gawne 2015; Hedinger 1984; H. Hill 1995; Larson 1978; McGregor 1994; Michael 2014; Munro et al. 2012; Nau 2006; Nikitina 2012b; Reid 1979; Rumsey 2010; Skjon 2001; Spronck 2015a; Widmer & Zemp 2017). While we agree with the analysis underlying alternative labels such as 'constructed' and 'represented' speech we believe that they may inadvertently suggest to refer to a different class of structures than the ones covered in the studies referenced above.

Most descriptive grammars that use the terms *reported speech*, *reported discourse*, *represented speech* or *quotation* appear to apply these notions to a crosslinguistically similar set of utterance types, seemingly without requiring a formal definition. However, in order to provide a rough idea of the type of phenomena we will discuss here in addition to the basic English examples in (1) and (2), we will set out by suggesting the very preliminary definition in (4):

(4) Reported speech is a representation of an utterance as spoken by some other speaker, or by the current speaker at a speech moment other than the current speech moment. For our current purposes, this includes all relevant meanings involved and the dedicated linguistic devices for signalling them.

Having proposed this definition (which we believe is compatible with most general discussions of reported speech in typological descriptions) we immediately need to acknowledge that this initial definition is problematic for three reasons: First, it explicitly applies to *speech*, i.e. spoken utterances. As in-depth studies of reported speech constructions, such as Larson (1978) and Güldemann (2008) have repeatedly demonstrated, however, in many languages the same type of representation applies to considerably more than just expressing speech, and a saying, i.e. an actual speech event need, not be implied by the construction, with meanings of, e.g. 'thinking' and 'wanting' being common as well.[3] While setting up a typology by starting from a prototypical use is not exceptional, we would prefer a definition of reported speech constructions that is more inclusive than the

2 The label 'reported discourse' perhaps even has a slight edge over 'reported speech' since it allows for other than spoken 'reports' (such as representing of thought or volition), but may also invoke a connotation of discourse as a 'stretch of speech', a text, which may less readily describe its status as a sentential construction.

3 Note that since Vološinov (1973) the term 'reported speech' has explicitly included non-speech functions, with the author paying much attention to 'inner speech', i.e. thought.

one in (4). Second, expressions of reported speech display great diversity, as we will also demonstrate in the following sections. Reported speech may be indicated with explicit structures that unambiguously mark the phenomenon, but also by minimal linguistic means, or even extra-linguistic means (such as eye gaze or gesture). As a consequence, the initial definition in (4) is at once too restrictive (when it comes to function) and too loose (when it comes to characterising reported speech as a linguistic category). A third problem with our preliminary definition in (4) is that it does not make a distinction between reported speech as a discourse act and as a grammatical construction. This allows us to be maximally inclusive with the initial set of structures and utterance types we can consider, but we recognise that as a formal definition (4) is ultimately unsatisfactory.

Before returning to the task of definition, however, we will first illustrate the breadth of the phenomenon by considering eight cross-linguistic features that, we claim, characterise reported speech in the languages of the world and set it apart from other syntactic categories.

2 Eight observations regarding the syntax of reported speech

The following sections introduce eight cross-linguistic observations about the syntax of reported speech that highlight distinct aspects of its formal expression and meaning across languages. These concern the syntactic relation between elements within reported speech expressions (Section 2.1), the fact that structural marking of reported speech is often extraordinarily variable, yet constrained (Section 2.2), the inherent semantic and pragmatic asymmetry[4] between the components of reported speech (Section 2.3), the occurrence of indexical elements in reported speech we do not find in non-reported speech related contexts (Section 2.4), striking processes of modalisation in reported speech (Section 2.5), the observation that the degree of shared grammatical values between the two components of reported speech appears to have predictable semantic effects without fundamentally altering the meaning of reported speech (Section 2.6), and the apparent regularity of diachronic and synchronic semantic changes in reported speech constructions, which suggest that they should be treated as a particular class of phenomena (Section 2.7). Finally, Section 2.8 notes that despite the fact that

4 We apply the term 'semantics' to any type of meaning that can be shown to be coded through a conventional form, and 'pragmatics' to any interpretation that can be related to general principles of usage, beyond individual constructions.

reported speech is an inherently conversational phenomenon, no language has been attested that cannot treat it as a dedicated sentential structure.

2.1 Reported speech constructions reflect a binary, but neither subordinating nor coordinating semantic structure (M:R)

Our first typological observation may seem relatively mundane, but relates to recognising reported speech constructions as a coherent class. Structures expressing reported speech typically consist of two separate morphosyntactic units (cf. Güldemann 2008), but the syntactic relation between these units is often rather idiosyncratic, both language internally and cross-linguistically. Consider the examples from Megeb Dargwa in (5)–(7) below.[5]

(5) Megeb Dargwa (Northeast Caucasian; Dagestan)
malla-rasbadi-j-ʔini ***ʔ-ib*****,** *ħa-la* *k'unk'ul-li-ʔini*
Molla-Nasreddin-O-ERG **say:PF-PST** you.sg-ERG cauldron-O-ERG
b-aq'-ib-il *k'unk'ur-g^{w}a* *iš.*
N-make:PF-PART cauldron-FOC this
'Molla Nasreddin said: "This is a cauldron born by your cauldron".'

(6) ***malla-rasbadi-j-ʔini*** ***ǯawab*** ***b-aq'-ib***, *ħa-la* *k'unk'ur*
Molla-Nasreddin-O-ERG **answer** **N-make:PF-PST** you-GEN cauldron
b-ebk'-ib ***ile***.
N-die:PF-PST **CIT**
'Molla Nasreddin answered: "Your cauldron died".'

5 None of the glosses in this paper have been adjusted from their original source. Glosses not in accordance with the Leipzig Glossing Rules are listed below, unless they are sufficiently self-explanatory (e.g. 'nonpast', 'past'). The distinction between capitals and lower case in glosses is not significant.

accm - accomplished form of the verb (Adioukrou); adr - addressee; AUG - augmented number; CARD - cardinal pronoun; CIT - citation marker (Megeb Dargwa); CM - conjugation marker (Nyulnyul); CONT - continuative aspect; CRD - cardinal pronoun; CSL - causal case (Arabanda); DUB - dubitative mood; EMP - emphasis; HPL - human plural (Megeb Dargwa); HS - hearsay marker; IMP - imperfective (Nyulnyul); imper - imperative mood; irreal - irrealis mood; LOG - logophoric; MIN - minimal number; npst - non-past tense; n_w - w-neuter gender (Ungarinyin); O - object; oblig - obligation; PART - participle; PERF - perfective; PF - perfective; PR - present tense; PRES - present tense; PURP - purposive; RED - reduplication; sp - speaker; SS - same subject (Usan); STAT - stative; succ - successive; UF - uncertain future (Usan).

(7) *ca* *adaj-la* *kʷi-jal* *urši* *le-b-re-kʷ-**an***
one father-GEN two-CARD son COP-HPL-PAST-**HS**
'A father had (**they say**) two sons.' (Ganenkov 2010: 1)[6]

Faced with the variety of morphosyntactic structures in (5)–(7), we may conclude that the relations between the bolded and underlined elements in these expressions of reported speech are radically diverse. These examples contain a complement construction with a simple verb in (5), one with a complex verb (*ǯawab b-aq'-ib* 'answer he-made') in (6) and a simple clause with a verbal inflection in (7). But this ignores an obvious semantic parallel between, e.g. all underlined clauses above, which describe an utterance the current speaker attributes to a reported speaker.[7] The elements in bold, either clausal or affixal, share a clear function as well: they index the speaker to whom the underlined clause is attributed. And as far as we are aware, the examples from Megeb Dargwa, and their English translations are representative in this respect of reported speech cross-linguistically. If we do not attribute these similarities to a supra-clausal level common to reported speech constructions, these typologically robust observations would be left unexplained.

The alternative view we would like to put forward here suggests that the similarity of the elements underlined and in bold in (5)–(7) follows from the analysis on which they form part of a dedicated syntactic construction, a reported speech construction. This approach goes back to an analysis first explicitly proposed by McGregor (1994), stating that reported speech necessarily involves a single type of syntactic relation, which he labels a 'framing relation'. This relation may hold between two clauses, or even between a modifying element and a clause. While paradigmatic oppositions between several types of reported speech constructions within a language (e.g. bi-clausal, as in (5) or (6), or morphological, as in (7)) yield functional differences between the constructions (Aikhenvald 2004:chapter 4; Spronck 2009), this approach allows us to classify these constructions as the same broad type.

As a shorthand, we will henceforth refer to the underlined clauses in (5)–(7) as R (for reported), and while R may be a full clause, the label may also apply to subclausal or multi-clausal structures. The non-underlined parts in (5)–(7), which show a greater morphosyntactic diversity than R, we call M (for matrix). M signals the identity of the reported speaker, which, as (7) demonstrates, may be a generic one.

6 We thank Mikhail Daniel for comments regarding the orthography of these examples.

7 Whether the reported speaker actually made this utterance is linguistically irrelevant, but the form of a reported speech construction often does signal whether the current speaker suggests that the utterance was made as such, or reflects an interpretation/reformulation by the current speaker (see also Section 2.6).

M and R may correspond to a bi-clausal morphosyntactic structure, but as (7) demonstrates, reported speech is not restricted to bi-clausal structures. M may be expressed as a morpheme, and can even remain unexpressed (see Section 2.2).

Postulating R and M as distinct semantic units allows us to draw parallels between reported speech expressions across morphological types and inter-sentential syntactic boundaries. As indicated, expressions of reported speech often employ a range of morpho-syntactic strategies, and our analysis allows for the classification of reported speech across these types of expression. Within this approach, expressions as in (5)–(7) can be recognised as members of a single constructional class. Positing M and R as semantic units that may correspond to structural units of various types allows us further to examine the formal and functional diversity found within reported speech. For example, by treating a hearsay morpheme as in (7) as M, we may shift our analytical focus to addressing the differences and correspondences between clausal, complement taking or morphological M, rather than assuming that, e.g. hearsay markers are only or primarily in semantic opposition with other members of an evidential paradigm. Several authors have argued that reportative and quotative evidential markers occupy a slightly idiosyncratic position within evidential systems (cf. Hengeveld & Hattnher 2015). Analysing 'saying'-clauses as M, rather than 'regular' main clauses or complement taking clauses also does more justice to cross-linguistic observations regarding these clausal constructions in reported speech: As McGregor (1994) points out, very few properties of, e.g. direct speech constructions in English qualify them as coordinated or regular appositional clauses, and indirect speech constructions as embedded or subordinate. For example, the cross-linguistically frequently attested flexibility in the positioning of matrix clauses with respect to R-expressing clauses, e.g. in interjecting position, is quite atypical of other types of coordination.[8]

2.2 Defenestration: Optional marking as a feature of reported speech

Despite some complex marking strategies found in reported speech (see, e.g. Section 2.4), a remarkably consistent observation about reported speech in conversation is that marking of M is often extremely limited or even seemingly completely absent (D'Arcy 2015). Any account claiming that reported speech requires analysis at a syntactic level needs to address cases in which the syntactic

8 For a fuller discussion of the distinctions between coordination and subordination on the one hand, and reported speech on the other, see McGregor (1994). See Spronck (2017: 106–107) for a summary of 14 arguments for why reported speech constructions are an idiosyncratic type.

relationship involved in reported speech is apparently not marked. Specifically, across languages meanings associated with M do not always receive structural expression. Consider, for example, the narrative passage in (8):

(8) Nyulnyul (Nyulnyulan; Australia)

a. *nga-walam-jirr / angki bur-uk / angki bur-uk /*
3NOM-PST-call:out-3AUG.OBL what place-LOC what place-LOC
nga-n-in /
1MIN.NOM-be-PRS
'He called out to them, "What country am I in?"'

b. *juy / mi-n-in / yarrad-nyirr / arri*
2MIN.CRD 2MIN.NOM-be-PRS 1AUG.CRD-COM not
mi-li-jid kalb /
2MIN.NOM-IRR-go up
'"You are here with us; you can't go back up there."'

c. *arri dumbar mi-li-j / in-uk / wamin /*
not fly 2MIN.NOM-IRR-say this-LOC different
yarrad-nyirr / judiny /
1AUG.CRD-COM straight
'"You can't fly away; you must stay here with us forever."'

d. *ngii / man / aa / angk-in / angk-in ka /*
yes but and who-ERG who-ERG cu
kad i-na-m-jan nhh / ni-marl /
cut 3NOM-CM-put-1MIN.OBL [laughter] 3MIN-hand
'"Yes, but who cut my arms?"'

e. *yarrad-mad-mad / kinyingk-mad / burrb / liyan /*
1AUG.CRD-EMP-EMP DEF-EMP dance like
mi-na-m akal /
2MIN.NOM-CM-put and
mi-n-di-jarrad /
2MIN.NOM-CM-say-1AUG.OBL
'"We did! You wanted it because you wanted to dance, so you told us!"'
(McGregor 2011: 716)

In line (8a) the common way of forming a reported speech construction in Nyulnyul is demonstrated, with a clausal M followed by R (McGregor 2011: 679ff). In the following lines, however, the M-clause is absent, even though the R is attributed to different speakers: the reported speaker of (8a) is the same as in (8d), an emu, who has found that his wings have been cut off, and the

reported speakers in lines (8b), (8c) and (8e) are his tormenters. (Note that the verb root glossed as 'say' may also act as an auxiliary/light verb in a so-called 'coverb', or 'preverb' construction and does not act as matrix verb in a reported speech construction in line (8a) (see McGregor 2014)).

Spronck (2017) proposes to interpret cases like the one in (8) as instances in which M-clauses are treated as optional elements, *sensu* McGregor (2013). In a pun on the label in McGregor (1994) for M, 'framing' clauses, we will refer to M-less clauses as 'defenestrated'. Crucially, the optionality of M in these instances is not semantic absence: if the situation requires it, a language could always add explicit M-marking, and to correctly interpret a defenestrated clause as in, e.g. (8) the semantic value of M (the reference to the emu as the reported speaker) has to be interpretable.

But in the absence of M, how can the hearer still identify the correct reported speaker in the defenestrated clauses in (8)? To this end, Nyulnyul uses a typologically common strategy: it constructs the reported exchange as question and answer pairs. In doing so, the illocutionary values of the subsequent sentences imply distinct speakers, rendering explicit marking of M redundant (see also section 2.2).

We propose that the defenestration analysis, i.e. an approach in which M may be treated as an *optional* element, can account for an observation about reported speech that would otherwise be quite startling: introductions to reported speech are extra-ordinarily varied. Consider the initial clause in (9):

(9) Arabana (Pama-Nyungan, Karnic; Australia)
Kadnhini *thurka-mda* *kudnala-ra:* *Wiyayi*, *kadnhiniya*
grandmother rise-PRES sleep-CSL boy.VOC grandson.VOC
intyamda *warra-nangka-rda?* *Padni-li.*
where play-CONT.S-PRES nothing-ADV
'The grandmother got up from her sleep: "Young boy! Grandson! where are you playing? (She called) in vain."' (Hercus 1994: 302–303)

Is the M element structurally marked in (9)? If our answer to this question is yes, we would have to conclude that the only element in which the reported speaker is explicitly indexed (i.e. the main function of M) is the initial clause *kadnhini thurkamda kudnalara* 'grandmother got up from her sleep'. This approach leads to a potentially infinite number of M types since there is no conventional restriction on the types of verbs and clauses that may occur before a reported speech construction. In our view, such an approach would obscure the cross-linguistic regularities that we actually find in reported speech. The defenestration analysis provides a principled alternative. The R element in (9) contains

several interactionally oriented elements: the two vocative nouns *wiyayi, kadnhiniya* 'young boy! grandson!' are unlikely to be addressed to the current addressee, the person listening to the narrative. The current addressee is also unlikely to be the target for the question asked in the following clause *intyamda warranangkarda?* 'Where are you playing?' An obvious candidate for a speaker deictically anchoring the vocatives and the question is the subject referent of the initial clause in (9) whose kinship description (*kadnhini* 'grandmother') also matches that of the reported addressee (*kadnhiniya* 'grandson'). All these clues suggesting a reported speech situation, with all relevant participants being deictically recoverable (not even taking into account potential prosodic cues indicating that (9) involves reported speech) mean that explicit marking of M is redundant. This does not mean that reported speech introducing clauses cannot show some cross-linguistic regularity,[9] simply that in (9) this clause is not the element in the syntactic relation that casts *wiyayi, kadnhiniya intyamda warranangkarda?* 'Young boy! Grandson! Where are you playing?' as R.

Our analysis of reported speech constructions as a syntactic class, combined with the proposal that M can remain an optional element in the expression of the syntactic relation involved further allows us to draw more fine-grained distinctions in the syntactic status of the corresponding constructional elements.[10]

2.3 Fundamental M-R asymmetry: M indexicality is pragmatic, R 'indexicality' primarily is not

The semantic and syntactic behaviour of M and R is remarkably distinct. One of these distinctions lies in the interpretation of deictics, more particularly, of

9 For example, they can be expected to describe relatively simple and predictable events without relative clauses ('grandmother awoke', as opposed to 'grandmother tried to balance a spoon on her finger while reciting a recipe for borscht, and skipping rope'), are likely past tense events concurrent with or immediately preceding the speech event and contain full lexical subjects.

10 The defenestration analysis also leads to a number of predictions. For example, in addition to more obvious interactional elements such as kinship terms, vocatives and interrogatives, R commonly hosts a number of word and indexical construction types that are less frequently found in other contexts, such as modals, discourse markers and (other) 'judge-dependent' elements (Stephenson 2007) or expressive vocabulary/ideophones (Dingemanse & Akita 2016; Nikitina 2012c) (also see Sections 2.3 and 2.5). These elements can be expected to be over-represented in defenestrated clauses as they are in R, as Spronck (2017) indeed shows to be the case for the Australian language Ungarinyin and Si & Spronck Forthcoming for the Dravidian language Solega.

pronouns (also see section 2.4). We posit the following cross-linguistic claim about the way in which deictic elements in M and R are interpreted:

(10) The interpretation of pronouns and other deictics in M is pragmatic, i.e. it relies on mechanisms in the *referential act* rather than strictly grammatical convention. For R the opposite holds: its interpretation is typically conventional and semantic, rather than pragmatic.

In order to illustrate the claim in (10) it may be most effective to start with the second qualification: elements in R are primarily conventionalised and conditioned by grammar in a narrower sense. This suggests that languages may show considerable variation with respect to what elements can undergo deictic shift and *how*, which Nikitina (2012a) shows to be the case. Languages vary, for example, in how they use deictic categories in reported speech constructions, and the rules for their use often depend on the specific type of construction (Comrie 1985: 302–303; Coulmas 1986b). In English, indirect reports differ from direct ones in the use of tense and temporal adverbials, cf. (11) and (12); in Russian, the two constructions do not differ in tense, cf. (13)–(14), and there are languages where the same deictic adverbials can be used as long as temporal reference is clear in context (Coulmas 1986b: 18, on Yoruba).

(11) John said yesterday, "I shall leave tomorrow".

(12) John said yesterday that he would leave today.

(13) Russian (Indo-European, Russia)
Včera Džon skazal, "Ja uedu zavtra."
yesterday J. said I go:NONPAST.1SG tomorrow
'Yesterday John said, "I shall leave tomorrow."'

(14) Russian (Indo-European, Russia)
Včera Džon skazal, čto uedet segodnja.
yesterday J. said that go:NONPAST.3SG today
'Yesterday John said that he would leave today.'

Similarly, languages vary in the way pronominal deixis works in reported speech constructions, and the most striking differences are dictated by the language-specific grammar rather than situational pragmatics or considerations of style (Nikitina 2012a). In English and other European languages, pronouns can receive different interpretations depending on whether they occur in direct

or indirect speech constructions (cf. the pronominal shift in examples (1) and (2)). In a number of African languages, third person pronouns can be interpreted, in the context of reported speech, as referring to the reported speaker, even though reported addressees must be referred to using 2nd person pronouns. For example in (15a) the second person plural *ɔny* occurs in R and is coreferential with *wɛl* 'them' in M. (By means of contrast, in Adioukrou, 1st person pronouns in R can only refer to the current speaker, and do not shift reference). In other languages, such as Gahuku (15b), 1st person pronouns are used to refer both to the current and to the reported speaker, i.e. in the context of reported speech they are inherently ambiguous:

(15) a. Adioukrou (Niger-Congo, Kwa; Ivory Coast)
li dad wɛl nɛnɛ ɔny *ùsr* **ir** *el*
3sg said.ACCM them this 2pl.REPORT build.IMPER 3sg.OBJ house
'She$_i$ said to them$_j$: "You$_j$ build me$_i$ a house (lit., you build her a house)."'
(J. Hill 1995: 91)

b. Gahuku (Trans-New Guinea, Gorokan Papua New Guinea)
geza ne-leqmo gilil-it-ove l-oko
you me-make heal-fut-1sg say-succ
'If you say you will make me healed (lit., I will heal me).' (Deibler 1976: 115)

In some languages, the idiosyncratic interpretation of deictic elements in the context of reported speech leads to characteristic patterns of agreement mismatches. In (16), the reported speaker is referred to by a 3rd person pronoun, but the verb carries a 1st person agreement marker. Such mismatches have been attested in reported speech constructions in various languages (for more examples and discussion, see Nikitina 2012a):

(16) Karimojong (Eastern Nilotic; Uganda)
àbù papà tolim ɛbè àlózì iŋèz *morotó*
AUX father say that 1SG.GO.NPST 3SG Moroto
'The father said that he was going to Moroto.' (Novelli (1985: 531), cited in Curnow (2002: 9))

These observations suggest that the interpretation of pronouns in R, as well as of other deictics, requires specific knowledge of the grammar of the language involved (also see Section 2.5). This phenomenon has been explored most extensively in formal semantics, particularly since Schlenker (2003), who challenges the following thesis which he attributes to Kaplan (1989): "The semantic value of an indexical is fixed solely by the context of the actual speech act,

and cannot be affected by any logical operators" (Schlenker 2003: 29). As indicated above, indexical elements in R are not in fact conditioned by the current speech situation, and, as Schlenker (2003) also points out, their value is language dependent. In Schlenker's (2003) influential proposal, this means that, *contra* Kaplan, sentences that involve 'attitude operators' such as reported speech constructions can introduce a local logical operator that 'fixes', for example, the person value as first person in (16). Kaplan (1989) calls such operators, which he claims not to exist, 'monsters', a label Schlenker (2003) and subsequent literature adopts to describe indexical effects in R (Deal 2017; LaTerza et al. 2015; Shklovsky & Sudo 2014).[11] While pragmatic factors in the speech situation may influence the behaviour of certain pronouns (Evans 2013; Maier 2017)[12] extra-grammatical factors do not determine the realisation of pronouns in R.

The literature on monsters demonstrates both the diversity of referential patterns in R and the consistency of these patterns in individual languages under a formal semantic approach. While details of the analysis may have different theory-specific implications, these observations suggest that indexical elements in R play a significant role in signalling reported speech.

Our claim for deictics in M is the polar opposite of our claim for R: pronouns in M refer 'normally' in the speech situation, as they do in regular non-reported speech utterances and in Kaplan's (1989) assumed default condition, and their interpretation can therefore be expected to be more homogeneous across languages. We believe that this analysis may be illustrated on the basis of another, initially puzzling observation: pronominal reference in M influences the degree of modal commitment towards the content of R. Specifically, the mention of a

11 As a reviewer points out, while Schlenker (2003) primarily associates monsters with specific indexical markers in the clause, such as speech verbs, many authors adopting the term apply it more broadly, even as a property of the entire construction. As in our proposal R is not a subordinate clause projected by a verbal head but a constructional element in its own right, only in this latter interpretation is the notion fully compatible with our proposal.

12 Evans (2013) notes that if the reported speaker is the current addressee, pronouns in R may exceptionally refer as second person subjects to the reported speaker, a phenomenon he calls 'second person magnetism'. For example, if John is present, in several languages it has been shown to be acceptable to say (the equivalent of) 'Marie said that you (=John) are here' rather than 'Marie said that John is here': the presence of the reported speaker 'attracts' the reference of the pronouns in R (but see Nikitina 2012a for an alternative interpretation). We believe that this is a typical instance of the exception proving the rule: this pattern is remarkable *exactly because* of the default referential patterns described here. For us, it also highlights the merits of treating R as a syntactic environment, rather than a purely semantic or pragmatic unit, because this allows us to distinguish between clearly pragmatic phenomena like second-person magnetism and more conventional semantic effects.

third person subject in M, as in (17), may often imply that the current speaker has doubts about the reliability of the reported message, whereas in (18), with a first person subject in M, the use of a reported speech construction rather strengthens commitment to the message that this is *really* my home.

(17) He said/was saying he lives here.

(18) I said/am saying I live here.

The first-person effect in (18) is not a quirk of English. For example, Michael (2010, 2014) cites examples of 'self-quotation' in the Arawakan language Nanti as in (19):

(19) Nanti (Arawakan; Perú)
noka hanta no=n-tim-e Kuriha-ku
1.quot there 1S=irreal-live-irreal Peach.Palm.Creek-loc
'I say, "I will live there at Peach Palm Creek."' (Michael 2010: 5)

First person subjects in M 'indicate a speaker's individuated commitment [to the content of R]' (Michael 2010: 5). In other words, in (19), the speaker is stating *hanta no=n-tim-e Kuriha-ku* "I will live there at Peach Palm Creek" as a very strong possibility, or sincere intention.

For Michael (2014), and several other authors who have replicated this observation in other languages since, the particular relevance of examples like (19) is that they seem to disprove a prominent claim in the literature about the semantics of reported speech, viz. the idea that one of its defining functions is to 'distance' oneself from a reported message:

> Reported discourse is the representation of a spoken or mental text from which the reporter distances him-/herself by indicating that it is produced by a source of consciousness in a pragmatic and deictic setting that is different from that of the immediate discourse. (Güldemann 2008: 6, Güldemann 2012: 118)

Whereas this general characterisation applies to non-contemporaneous (i.e. typically past tense) and non-current speaker originated (i.e. typically third person) reported speech, as in (17) and the majority of examples Güldemann (2008) cites, it does not apply to reported speech with first person subjects in M, as in (18) and (19).

We agree with Michael (2014) that the observation that reported speech may both involve strengthening and weakening commitment to the content of R suggests that 'distancing' is perhaps not the right label or even the

appropriate defining component of a typological definition of reported speech (see Section 3). Note, however, that the commitment properties of (18) and (19) can easily be related to more generally observed person effects, as, e.g. captured by Silverstein's (1976) referential hierarchy. Exceptional cases aside, such as, e.g. third person referents that index highly authoritative people, which may thereby typically count as a reliable source, first person subjects in M typically enhance commitment, third-person reference typically decreases it. A second person M (e.g. 'you said you would do the dishes today') suggests an appeal of commitment by the addressee. The mechanism through which these meanings seem to arise is pragmatic implicature (cf. Michael 2014: 187).

In this sense we observe a fundamental asymmetry between the semantic and pragmatic status of M and R: in R, indexicals are primarily conditioned by grammatical rules that are specific to R as a unit within the reported speech construction. In M, we can explain the observations on the basis of more general principles of person reference that are not primarily specific to M, a difference we label 'pragmatic'.

Given the connection between the interpretation of M with more general principles of language use, we suggest that this interpretation is more homogenous across languages. In any case, to our knowledge no instance of reported speech has been described so far in which similar degrees of variation found with respect to the referential patterns of R, are also observed in M. These observations point to a fundamental asymmetry between M and R.

With Güldemann (2008) we do agree, however, that the commitment effects observed in reported speech (although not only distancing, but also strengthening commitment) need to play a role in the definition of reported speech, and we will explicitly attempt to acknowledge this in the definition of reported speech constructions we propose in Section 3.

2.4 R involves specialised indexical markers

In 2.3 we discussed ways in which pronouns may receive interpretations in R that are not available to them in other contexts. It is also quite common to find in R pronouns that do not occur outside constructions with reported speech. An example of this is the use of specialised logophoric markers in African languages (Clements 1975; Hagège 1974; Nikitina 2012b, Nikitina 2018), and also beyond the African continent (Bugaeva 2008; Daniel 2015; Demirok & Öztürk 2015; Nau 2006). In (20), a special logophoric pronoun is used to refer to the reported speaker; in (21), special logophoric pronouns are used to refer

both to the reported speaker and to the reported listener. As the example in (20) shows, the use of logophoric pronouns cannot be explained in terms of subordination, as they also appear in constructions that are clearly not embedded (for a full argument, see Nikitina 2012b).

(20) Wan (Niger-Congo, Mande; Ivory Coast)
ké lā zò-á bā biàgà nɛ̀ zɛ̄ zānā̠ dì
if 2SG come-STAT.PERF LOG.SG wake PURP word true say
[He said] '"If you've come to wake me up, tell the truth."' (Nikitina 2012b: 289)

(21) Goemai (Afro-Asiatic, Chadic; Nigeria)
k'wal yin gwa *goe tu* ji
talk say SG.M.LOG.ADR OBLIG kill SG.M.LOG.SP
'(He_i) said that he_j should kill him_i.' or 'He said: "You should kill me".'
(Hellwig 2006: 219)

Languages vary in the details of how such specialised markers are used, and in the range of contexts where they appear, yet the underlying phenomenon is the same: speakers choose from different inventories of indexical markers in R and in other contexts. The existence of logophoric pronouns further supports the validity of recognising R as a distinct structural unit.

2.5 Modal shift in reported speech involves bi-perspectivisation

Deictic pronouns are not the only type of element whose behaviour is affected by their involvement in reported speech constructions. When an epistemic modal marker is used in reported speech, it can also adopt a specific meaning it does not have outside reported speech, and, seemingly this may affect elements in both M and R. Furthermore, the change in meaning is not random, but typically involves a shift towards a 'multiple-perspective' (Evans 2006) or 'bi-perspectival' meaning (Evans 2013).

Consider examples (22), (23) and (24) in which the authors interpret markers that in other contexts receive a regular epistemic modal meaning, with a more complex type of epistemic evaluation roughly along the following lines: 'reported speaker believes/says that *p*, but I (now) state that not-*p*'.

(22) Western Tarahumara (Uto-Aztecan; Mexico)
simi-le-ga-ra-e
go-PAST-STAT-QUOT-DUB
'Someone said he went but he did not.' (Burgess 1984: 104)

(23) Ungarinyin (Worrorran; Australia)
goanna-karra nga-ma-ra nya-langkun kuno
goanna-maybe 1sg-do-PST 3fsg-head n_w-DIST
'I thought it was a goanna's head over there [but I turned out to be wrong].' (Spronck 2015b: 178)

(24) Aguaruna (Jivaroan; Perú)
Tikima maanchuchi. Wainchauk "Uchuchiji," tumalnai.
very it-is-small one-who-has-not-seen-it it-is-his-child he-might-say
'It is very small. A person who had not seen one might believe (incorrectly) that it was a baby (bird)' (Larson 1978: 107)

All three examples (22)–(24) have the form of a reported speech construction: (22) has a quotative marker, in (23) *ngamara* lit. 'I did' represents a typical M-form in Ungarinyin and can equally be translated as 'I said that...' or 'I thought that...' and in (24) M is formed by the word *tumalnai* 'he might say'. In addition, each example includes an epistemic modal marker, glossed as dubitative (22), 'maybe' (23), and 'might' (24), respectively. However, the resulting interpretation is *not* dubitative: it is as indicated above, 'x said/thought that *p*, but I (now) know that *p* is wrong'. These 'mistaken belief' expressions require a fuller typological treatment than we can provide here, and we will not speculate about the precise mechanisms underlying the meanings of (22)–(24) (but see Section 4; also cf. Spronck 2015b).[13] This suggests specific constructional effects between elements at the level of the entire reported speech construction, not just M or R. The modal marker that gives rise to the mistaken belief meaning in (24) supports this assessment as well, since it appears in M, suggesting that bi-perspectivisation may occur throughout the reported speech construction, rather than either in M or R.

13 One reviewer suggests that understanding scope relations is central to understanding how these interpretations arise. For example, if *-karra* 'maybe' in (23) only has scope over M this may lead to another interpretation than when it has scope over the entire sentential construction. In our estimation, the epistemic modal marker in each of the examples (22)–(24) has maximal scope, i.e. encompasses the entire reported speech construction. These scope properties alone would not yet explain, however, why these examples do not just mean '*x* maybe said *p*', but receive the interpretation '*x* incorrectly believed/maintained that *p*'.

A mistaken belief interpretation is a bi-perspectival meaning: it indexes the perspective of both the current speaker at the current speech moment, and that of a reported speaker. The observation that such complex perspectival meanings arise in reported speech constructions, seems to indicate a pattern in reported speech constructions that further adds to their complexity and idiosyncrasy. Bi-perspectival meanings are not restricted to constructions with modal markers either. For example, Widmer and Zemp (2017) observe that in Tibetan languages, pronominal markers in R have adopted bi-perspectival epistemic meanings. Specifically, in examples such as (25) the copula does not agree with the first person subject in R, since it is marked as 'allophoric' rather than 'egophoric'.

(25) Shigatse Tibetan (Sino-Tibetan; Tibet Autonomous Region)
kʰœ̀ ŋa̱ pʰœ̱pā pię̱ sa̱
3sg.ERG 1sg Tibetan COP.ALLO say.PFV
'S/he said (that) I am Tibetan.' (Haller & Haller (2007), cited from Widmer & Zemp (2017: 57))

In Widmer & Zemp's (2017: 61–62) formulation, instances of pronominal inflection as in (25) 'cease to bear a syntactic relation to the person subject of a clause and begin to bear an epistemic relation to the assertor of the proposition, [...] that is the reported speaker.'

We suggest that these examples indicate once more that reported speech constructions constitute a distinct class, characterised by a particular proneness to bi-perspectival meanings. We will also propose that these complex perspectival meanings have to be an integral part of their definition (see Section 4).

2.6 Grammatical choices in reported speech constructions reflect perspective

The interpretation of modal markers in reported speech in Section 2.5 introduced the notion of 'perspective', which, we argue, is also central to the analysis of another phenomenon commonly distinguished in classifications of reported speech: the opposition between direct and indirect speech (i.e. constructions as in (1) and (2), respectively). Evans (2013) argues that this distinction is not binary, but scalar and that coding choices along this scale reflect calibrations of *perspective*.

The traditional standard semantic interpretation of this opposition has long remained as it was already presented in prescriptive grammars of Latin, along the following lines:

> [a] Direct Quotation [i.e. direct speech] gives the exact words of the original speaker or writer (*Ōrātiō Rēcta*). An Indirect Quotation [i.e. indirect speech] adapts the words of the speaker or writer to the construction of the sentence in which they are quoted (*Ōrātiō Oblīqua*). (Greenough et al. 1903: 374)

It is unlikely that this distinction ever represented an actual characterisation of the use of 'direct' and 'indirect speech' in a descriptive sense. Classicists have found repeatedly that when considering actual language data in Latin or Greek, the distinction between the two sentence types is rarely fully discrete or implemented as prescribed (Andrewes 1951; Gildersleeve 1906; Postgate 1905; Salmon 1931). The standard interpretation of direct speech comes close to the meaning of legalistic phrases such as 'and I quote', implying maximum impartiality on the part of the reporting speaker, but this does not mean that a reported message mechanically reproduces 'the words of the original speaker' even in direct speech. Both direct and indirect speech are necessarily imprecise reproductions of the reported utterance, if only because even the most skilled imitator could not switch to the literal acoustic voice of the reported speaker (Plank 1986: 285), or simply because of imperfections of memory. From a linguistic perspective, attempting to characterise semantic distinctions in reported speech on the basis of whether there exists an actual 'real-world' utterance of this exact same form or not is a fruitless exercise (see Vandelanotte 2009: 118–130 for discussion, and an extensive list of references finding 'non-verbatim' examples of direct speech).

In addition to the problematic traditional semantic opposition between direct speech and indirect speech, identifying robust formal distinctions between the two, both intra- and cross-linguistically, has proven challenging (Coulmas 1986b). The earliest example of a reported speech phenomenon that resists binary classification as either direct or indirect speech is free indirect speech (Bally 1912; Lips 1926; Vološinov 1930). In its most simple form this phenomenon corresponds to a reported message without a matrix clause,[14] like the second clause in (26), which Banfield (1982: 291) cites from Jane Austen's novel *Emma*.

(26) Never had she felt so agitated [...] She felt it at her heart. How could she have been so brutal, so cruel to Miss Bates!

14 As indicated in Section 2.2, we consider unexpressed matrix clauses a symptom of the syntactic optionality in reported speech constructions, which means that some forms of free indirect speech can be analysed as a type of variability in marking. This does not preclude the possibility that languages may also distinguish free indirect speech as a separate construction type contrasting with other types of reported speech constructions.

The final sentence in (26) represents an internal monologue on the part of the protagonist introduced as 'she' in the first clause above. Bally (1912) distinguishes no less than four different types of free indirect speech, and Lips (1926) provides an extensive overview of the many labels that have been given to semi-(in)direct reported speech phenomena falling between or beyond the binary opposition of direct and indirect speech. The discovery of logophoric pronouns (Hagège 1974), as in (11), complicated the formal distinctions even further, since in logophoric constructions most deictics refer as in the reported speech situation (which implies direct speech), yet the logophoric pronouns are a feature of the reported speech construction, which implies an adaptation to 'the words of the (current) speaker', and, hence, indirect speech. De Roeck (1994) finds that languages, in fact, vary widely in the type of deictic features that are 'adjusted' to the current speech situation, and the degree to which this occurs. These observations, and many similar ones made more recently, suggest that the opposition between direct and indirect speech – whatever it is still worth – is not a binary distinction, but involves a range of intermediate types (Aikhenvald 2008; Evans 2013; Maisak & Merdanova 2010).

The re-evaluation of direct/indirect speech, from discrete categories to a more scalar phenomenon, is consistent with Evans' (2013) perspectival interpretation of the direct-indirect speech contrast. Along the scale from direct to 'canonical indirect' speech, not only does 'direct speech represent a greater degree of distancing of the proposition from the speaker than indirect speech' (McGregor 1994: 82), it suggests a mix of perspectives in which that of the reported speaker is dominant. Within this view, 'less' direct, semi-direct, or 'indirect' speech constructions reflect an increasingly greater contribution of the perspective of the current speaker. Within this analysis, 'mixing' perspectives is the norm, and a central property of reported speech, a view that ultimately goes back to the claim by Vološinov (1973) that reported speech constructions are necessarily 'multi-voiced'.

The relevance of the labels 'direct' and 'indirect' speech is particularly sensitive to cross-linguistic variation. This not only applies to the shade of 'indirect' speech we observe. Rumsey (1990) rightly points out, for example, that despite the frequent claim that direct speech is universal, a direct speech construction in a language that lacks a contrast between direct and indirect speech cannot carry the same semantic load as in a language that *has* the opposition (or even multiple contrasts between reported speech constructions; also see Nikitina 2012b). However, for the argument we aim to develop here the main observation is that in reported speech constructions the distinction between syntactic choices has direct effects on to what extent the perspective of the current speaker coincides with that of the reported message. As far as we

are aware, this is a phenomenon not attested for any construction type other than reported speech, and therefore testifies to its uniqueness as a syntactic construction.

2.7 M-shift follows an evidential-modal-aspectual cline

Reported speech constructions appear to develop a remarkable range of polysemy in the languages of the world (Chappel 2012; Güldemann 2008; Larson 1978; Matic & Pakendorf 2013; Pascual 2014; Reesink 1993; Rumsey 1990; Saxena 1988; van der Voort 2002).

The multifunctionality of reported speech both involves grammatical functions and a number of common lexical changes in reported speech constructions that have been attested cross-linguistically.[15] These broadly deal with evidential, modal, and aspectual meanings, as we will illustrate with (27)–(30).

(27) Beijing Mandarin (Sino-Tibetan; China)
我 总是 觉得 说， 生活里。 缺了 点儿 什么
wǒ zǒngshi juéde shuō, shēnghuó-lǐ quē-le diǎnr shénme
1SG always feel say$_{THAT}$ life-in lack-PFV little something
'I've always felt that there is something a little lacking in my life.' (Fang Mei, in Chappel 2012: 84)

(28) Maybrat (West Papuan; Papua New Guinea)
y-awe y-aut ara
3M-say 3M-climb tree
'He says he climbs into the tree'/'He wants to climb into the tree.' (Dol 2007: 78)

(29) Usan (Trans New Guinea, Papua New Guinea)
Mi qei-qei mani umer-iner qamb gitab ig-oun
thing some-RED yam wilt-3s:UF say:SS abstain:SS be-1p:PR
'We abstain from various things lest the yam wilts.' (Reesink 1993: 222)

15 An interesting observation in this context is that Heine and Kuteva (2002: 261–269) list no less than eight functions the verb SAY can potentially grammaticalise into, almost the maximum number of functions of any entry in their dictionary of grammaticalisation. We would hypothesise that this is a consequence of the variety of syntactic frames the verb SAY may combine with, but, given their lexical orientation, Heine and Kuteva (2002) provide insufficient illustration to verify this.

(30) Wan (Niger-Congo, Mande; Ivory Coast)
yī ē gé ɓā kɔ́
water DEF say LOG boil
'The water started to boil.' (lit., 'the water said: "let me boil"')

In the Beijing Mandarin example (27) the verb 'say' is used as a light verb and the main semantic content of the sentence has the form of a speech complement, but the full construction is translated as an internal sensation. In (28), we find an even clearer example of reported speech, as indicated by the first translation given, but an equally adequate translation interprets the reported message as a volition. In the Usan example (29) a yam is seemingly presented as part of a reported speech expression stating 'we say that the yam might wilt', but the interpretation is an apprehensive clause, i.e. the reported message is an undesirable possibility. In (30), as the literal translation indicates, the reported message 'let me boil' is attributed to the water, but the resulting interpretation is an aspect-like one, in which the beginning of the boiling process is highlighted.

Although examples like these have now been described for a wide range of languages (see the references in the first paragraph of this section), the attested polysemies deserve a fuller typological analysis than we are able to provide here. However, there is a striking aspect to the non-speech interpretations found in these reported speech constructions. While they initially seem wide ranging, the meanings found so far do not point to the conclusion that they form an unrestricted set. Impressionistically, they appear to form an implicational scale in the languages that we surveyed, such that languages that allow an aspectual interpretation of a formal reported speech construction, should also allow modal and evidential interpretations. Güldemann (2008), who describes an even broader range of secondary meanings for markers of reported speech in African languages, cites the following implicational scale: *quote* > *complement* > *purpose* > *reason* and/or *condition* > *other* (Güldemann 2008: 523).

The main conclusions we would like to draw from these observations are, first, that reported speech constructions can serve as source constructions for a variety of other functions. This observation supports our view that they constitute a specific set of syntactic units, rather than, e.g. just regular complement clauses, because we do not observe this type of constructionalisation with complement clauses across the board. Second, and more speculatively, the observation that reported speech appears to give rise to a range of functions that bear no obvious relation to 'saying' or 'thinking', but appear to have cross-

linguistic similarities, suggests that reported speech constructions can be associated with a complex of semantic functions that may emerge in diachronic constructionalisation. In Section 3.3 we suggest a few properties that may be semantically strengthened or bleached in the course of this diachronic process, in addition to the ones proposed by Güldemann (2008). For our present purposes, however, the observations in this section mainly serve to further underline the idiosyncrasy of reported speech as a constructional unit.

2.8 Reported speech constructions are conditioned by grammar

Our final observation concerns a property of reported speech that has received a great amount of attention, particularly in the recent literature: reported speech appears to be accompanied disproportionally by gesture and prosodic cues.

Although claims about the contribution of such signals are rarely quantified, they have given rise to a view in which reported speech is a 'multimodal' phenomenon, which is not signalled through language structure *per se*, but is equally created through special gestures and voice quality (Blackwell et al. 2015; Dancygier & Sweetser 2012; Lampert 2013; Stec et al. 2015). The implication of these accounts is that morphosyntactic marking is just one of a range of linguistic and extra-linguistic signals that contribute to the expression of reported speech. D'Arcy (2015) goes even so far as to deny syntax any special privilege in the description of reported speech at all. 'Multimodal' analyses are further prompted by observations as discussed for the Nyulnyul example in (8), showing that in many languages it is possible to report speech by mimicking taking turns between, e.g. an enacted version of the current speaker at an earlier speech moment and a reported speaker.

As Stec et al. (2015) demonstrate, however, in one of the few detailed studies of multimodality in reported speech, no multimodal signal in reported speech plays a conventional role that compares to morphosyntactic expression (also cf. Malibert & Vanhove 2015). While the expressive nature of reported speech stands out, gesture or prosody have not yet been convincingly demonstrated to contribute to its *marking*. Based on the available evidence, we therefore find it improbable that multimodal strategies can have the grammatical status of a marking element in a language L, but will remain agnostic on the possibility that they may have such a status in some languages. In any case, reported speech is accompanied by multimodal properties to a striking degree, and these need to be given a proper and principled place in its description.

3 A proposal: Defining reported speech as a cross-linguistic syntactic category

3.1 A definition of reported speech constructions

The definition of reported speech constructions we would like to propose is the one in (31):

(31) A reported speech construction minimally includes three meaning components:
   a. A semiotic status of 'demonstratedness'
      This component qualifies the reported message as a representation of an utterance, as *demonstrated* discourse;
   b. Evidentiality
      This component reflects a deictic relation between two events: the alleged original situation of discourse production (the 'source of information') and the current speech moment;
   c. Modality
      This component reflects an epistemic evaluation of the represented utterance, allowing the reported speaker to qualify the reported message, either by strengthening or weakening truth commitment.

Our claim is that in order to be recognised as a reported speech construction, at least all the elements in (31) have to be directly relevant for the interpretation of a construction C. Therefore, for every reported speech construction it should be possible to identify each of the meaning components in (31).

In terms of the simple English examples in (1) and (2), the components in (31) may be identified as follows: The semiotic status of 'demonstratedness' as intended in (31) applies to the R-clauses 'Look, there is marmalade here' in (1) and 'there was a typhoon yesterday' in (2). Semiotically, these clauses do not reflect simple 'symbolic' language in the sense that they are arbitrary signs that stand for some experience the speaker shares with her addressee. They are 'intended to depict (mimic, simulate, provide an iconic representation of) the target of quotation: [the reported speaker]'s utterance' (Recanati 2001: 642). Roughly, this property has been captured in terms of 'demonstration' (Clark & Gerrig 1990), or 'depiction' (Clark 2016), and, more recently, it has been called 'iconic' (Davidson 2015; De Brabanter 2017; Recanati 2001; Spronck 2017). While these accounts may differ on details, the basic insight captured by these terms is

relatively straightforward: the current speaker presents the reported part of the reported speech construction as a typified version of an utterance she is attributing to a reported speaker, i.e. she signals 'this is a stylised version of an utterance'.[16] As a shorthand, we will label this property 'semiotic status as demonstrated'.[17] The term 'demonstrated', rather than 'demonstration' used by Clark and Gerrig (1990) underlines that we analyse the semiotic status of the reported element as a result of a demonstration, instead of a process of 'demonstrating' or 'depicting'. This interpretation is also captured by the term 'icon' in, e.g. Recanati (2001). Clark & Gerrig's (1990) notion of 'demonstration' as a process or act does suggest a second property, however, which 'demonstrated' does not: the demonstrated part of a reported speech construction is interpreted against the background of a broader event, more particularly two events, namely the event in which the demonstrated element is understood to have occurred, and the event in which it is demonstrated (i.e. the current speech moment). We would like to acknowledge this specific meaning as a separate meaning component however, which we label the evidential meaning of the reported speech construction.

The evidential meaning sets up a deictic relation between the demonstrated element and an alleged reported speech event, i.e. the construction 'points' to, or indexes, a reported utterance, or more specifically, the event in which the reported utterance was allegedly perceived (the perception event, or 'source of information'). This indexical meaning is inherent to the semantics of evidentiality (cf. Haßler 2002, 2010) and goes back to its early definition in Jakobson (1957), who defines it as consisting of three events representing (1) an event/proposition that is being talked about, (2) an event in which the proposition was allegedly observed, which we may call 'perception event', and (3) a current speech event.[18] Impressionistically, the Jakobsonian definition of evidentiality

16 The observation that reported speech constructions involving markers such as 'like' and 'type' occur across (European) languages (Davidse et al. 2013; Kolyaseva 2018; Sergeeva 2010; Wiemer 2010) may add further support for this analysis.

17 In formal semantic terms, this property is also related to the distinction between *using* words and *mentioning* them (cf. Davidson 1979). A 'demonstrated' element is both used (i.e. in order to describe something) and mentioned (a word as-such).

18 Jakobson (1957) refers to the first two event types as 'narrated event' and 'narrated speech event', respectively. The three-way event classification may be formalised as '$E_{proposition}$ E_{source} / E_{t0}', with the first two events carrying a deictic relation with respect to the current speech event (the interpretation of $E_{proposition}$ E_{source} is normally understood as not-present, not-here), which Jakobson (1957) indicates with the forward slash '/' between the first two events and the current speech event (E_{t0}). The deictic relation between the event types on both sides of the forward slash corresponds to the two-way distinction between the event in which a demonstrated element is understood to have occurred, and the event in which it is demonstrated as described

may seem distinct from more familiar definitions of the category as 'source of information' (Aikhenvald 2004), but note that they fundamentally share the same analysis: the function of evidentiality is to present a proposition as having been perceived in some way (in reported speech specifically through conversation). Aikhenvald's (2004) definition highlights the perception event, whereas Jakobson (1957) deconstructs the evidential meaning slightly further (and focuses on the relation between the events involved). However, the two definitions naturally imply each other.

With respect to the English examples in (1) and (2) at the beginning of this paper, this means that the clauses 'Look, there is marmalade here' and 'there was a typhoon yesterday' are marked by the construction as propositions/content (Jakobson's first event type) situated in a source or perception event (i.e. the second event type), and the phrases 'John said' in (1) and 'John says that' in (2) reflect the relation between these two events and the current speech event. The latter part of the construction most clearly brings out the deictic nature of the relation between the event types by the use of tense, with the past tense matrix verb in (1) placing the first two events in the past relative to the speech moment, and the present tense signalling the continued relevance of the first two events at the speech moment in (2).

Both the semiotic status of the reported utterance as demonstrated and the status of the relation between the events of the evidential meaning are subject to considerable cross-linguistic variation, depending on the language-internal opposition between constructions available in specific languages. But the discussion about 'distancing' and bi-perspectivisation in Section 2.5 suggested that for a full semantic characterisation of reported speech a third semantic component needs to be taken into account: a modal meaning, more particularly, an epistemic evaluation of the proposition as either credible or not.

While our notion of modality is broader than the 'distancing' that was central to Güldemann's (2008, 2012) definition of reported discourse cited in Section 2.3, it builds on the same analysis: reported speech typically involves an element of suspension of belief, or not committing to the truth of the reported proposition (Spronck 2012).[19] In other words, in (1) and (2) the current speaker

above. For further discussion of the terminology and formalization of Jakobson's (1957) definition of evidentiality, see Spronck (2015c), and references therein. Note that the perception event corresponds to the 'source of information' in more commonly cited definitions of evidentiality (cf. Aikhenvald 2004).

19 We believe that our definition in (31) is compatible with the definition by Güldemann (2008, 2012), since the notion of 'representation of a spoken or mental text' relates to a similar property as our semiotic status of the reported element as a demonstration, and the 'deictic setting that is different from that of the immediate discourse' is captured by the default evidential meaning.

does not only signal that John made statements about marmalade and typhoons, but also that she cannot vouch for these statements. As shown by examples such as (22), (23) and (24) in Section 2.5, reported speech may also allow for more complex types of current speaker evaluations. Although across languages, evidential and epistemic modal meanings are frequently found in the same contexts, a fundamental characterisation of the semantics of both evidentiality and epistemic modality requires them to be defined separately, if only to understand their various interactions (Aikhenvald 2004: 5; Cornillie 2009). We propose that a reported speech construction as a syntactic unit in a language L represents a specific calibration of evidential and modal meanings, as well as a conventionalised indication of the semiotic status of the reported element.

3.2 Using the definition

In order to serve as an explanatory and diagnostic tool, our definition needs to be able to (1) distinguish between reported speech constructions and non-reported speech constructions, and (2) explain the observed idiosyncrasies of reported speech signalled in Section 2. We dedicate Section 3.3 to the second task, but will first attempt to show in some detail how the definition in (31) allows us to identify a reported speech construction.

A useful place to start this exercise may be to examine potential 'borderline' cases of reported speech as in (32)–(35).

(32) He believes that there is no tooth fairy.

In (32) at least two interpretations are available. The sentence could simply represent a description of a belief that the speaker thinks the referent holds; under this interpretation it would not seem accurate to us to describe (32) as reported speech.[20] However, the example could also be used to represent the

The definition in (31) presents these meaning components in a more compartmentalised fashion, however, which, we think, makes it a more suitable starting point for typological comparison.

20 As one reviewer remarks, 'believing that not *p*' is a rather atypical belief sentence, but we use the phrase 'holding a belief' here as a property that can be held over both positive and negative propositions. Within this interpretation the sentence 'He believes that there is a tooth fairy' illustrates the same belief type in (32).

utterance "there is no tooth fairy," in which case it could be characterised as reported speech under the inclusive definition we suggested in (4).

While we argued that whether an utterance U portrayed in reported speech *actually* existed as such in the real world is linguistically irrelevant (see Section 2.6), the *interpretation* whether (32) reflects an utterance or not has direct influences on, e.g. what elements may occur in the complement clause in (32), how it should be paraphrased and what prosodic patterns are expected. For example, in (32') the inclusion of an interjection or an element such as '*na-na-nah-na-naah-naah*' is only possible if the entire sentence is understood as reported speech. Example (32") would probably be judged a correct paraphrase of (32) by most speakers of English, but is only possible under a reported speech interpretation. Finally, the prosodic break in (32"') is only consistent with an interpretation in which "there is no tooth fairy" should be attributed to the referent as an utterance (or constructed utterance, in the sense of Tannen 2007).

(32') He believes that, *na-na-nah-na-naah-naah*, there is no tooth fairy.

(32") According to him there is no tooth fairy.

(32"') He believes that [prosodic break, change of voice to high pitch] there is no tooth fairy.

We will further discuss how to weigh the evidence presented by (32')–(32"') shortly, but the main point these observations intend to illustrate here is that the semantic choice whether these examples constitute reported speech affects the semantic status of (31) *as a linguistic unit*. As soon as the complement clause is understood as a reported message, the main clause is understood as describing a reported speech event. In other words, as soon as we understand 'that there is no tooth fairy' as R, the element 'he believes that' *has to be* interpreted as M. In our analysis, this would qualify (32) as reported speech, at least under one of its interpretations.

The main contribution of our definition in (31) is that it allows us to distinguish between the reported speech and non-reported speech interpretations of (32). If (32) is reported speech, the phrase 'that there is no tooth fairy' has to be interpreted as 'demonstrated' in the sense introduced in Section 3.1. Significantly, it also needs to have an evidential meaning, i.e. 'that there is no tooth fairy' needs to be understood as observed in a perception event (i.e. a reported speech event) and, it needs to allow for a commitment effect, i.e. a modal distancing (or truth enhancing) interpretation. Our claim is that for a reported speech interpretation,

all of these three meaning components need to be present, if one or more lack in the interpretation, (32) is not reported speech.

This case could perhaps be made even more explicitly on the basis of (33).

(33) He sees that she is entering the room.

Example (33) is a factive sentence (Gentens 2016; Kiparsky & Kiparsky 1970) in the sense that adding a modal qualification of the kind 'but I doubt that' seems slightly odd. The phrase 'that she is entering the room' is also not commonly interpreted as a demonstrated element, nor as mediated through a perception event. Therefore, none of the three meaning components of our definition in (31) seem to apply, classifying (33) not as reported speech. But now consider (33').

(33') He sees that –ah!– she is entering the room!

Like all other English examples in this paper, (33') is constructed, and it could be debated whether (33') is a grammatically well-formed utterance or would be common in a description of English discourse. For our purposes, however, it illustrates a crucial property: if (33') is well-formed, the only way it can be interpreted is under a reading in which 'she is entering the room' is either a thought or utterance attributed to the referent of 'sees'. In other words, the addition of 'ah!' is only interpretable if the entire sentence is understood as reported speech. Again, this is a categorical decision that applies not just to some parts, but to the entire sentence in (33').

How to interpret this observation? As some of the authors cited in Section 2.8 have claimed, we could say that in sentences such as those in (32'), (32"') and (33') the elements 'na-na-nah-na-naah-naah', the pitch reset and 'ah!' *make* the utterance a reported speech construction. Such an assessment may not seem unreasonable given that these elements provide the only formal contrast with (32) and (33), which are not (necessarily) reported speech. A problem with this analysis, however, is that the empirical distribution of interjections in reported speech (cf. Spronck 2017) or of prosodic cues (Malibert & Vanhove 2015) does not seem to systematically mark reported speech, at least in the languages examined so far. If these strategies indeed have the status of linguistic *markers*, such findings are unexplained.

A more principled account, we suggest, would be based on a view in which reported speech constructions provide a receptive syntactic environment for elements such as interjections and expressive prosody, but are not determined by these. Under this assumption, such elements, and perhaps also other multimodal signals, would be more frequent in reported speech, but would not count

as markers of the category. Such a view can only be substantiated, however, if we provide an independent definition of what counts as reported speech, and what does not. And this is exactly what (31) aims to provide.

In example (33'), 'she is entering the room' is a demonstrated element, not a description of some world, but a proposition that was *perceived*, i.e. it needs to carry an evidential meaning. The addition of an epistemic qualification such as 'but I doubt that' seems marginally more acceptable for this example. Attributing the element 'she is entering the room' to a referent as an utterance or a thought means that the current speaker may agree or disagree with this phrase, without contradicting the meaning of the sentential unit in (33').

For performative utterances such as (34), the decision whether they constitute reported speech or not depends on how we value the demonstrated and evidential meanings.

(34) I am telling you that he is in for a surprise.

The commitment enhancing meaning of (34) is fully consistent with the observations about how person reference in M affects pragmatic judgements about R, but since we have also called this a more general pragmatic principle it does not necessarily derive from the inherent modal meaning of reported speech. The semiotic status of 'he is in for a surprise' is more questionable: although it is made somewhat prominent by the clause 'I am telling you', the phrase is not demonstrated in the sense that it is a stylised version of an utterance, the current speaker actually means to *do* the utterance 'he is in for a surprise'. With respect to the evidential meaning, we suggest that the event structure, with a perception event somehow contrasting or interacting with the current speech event is irrelevant for (34). In our analysis, these judgements mean that the semiotic and evidential meaning components required for reported speech are absent in (34), and that this sentence therefore represents a different syntactic class (for example, that of explicit performatives). This corresponds to our intuitive assessment that (34) is indeed not an example of reported speech, and the definition in (31) provides criteria for validating this assessment. The analyses of what counts as a demonstrated or evidential meaning could be further contested by those who would want to include (34) among the class of reported speech, however, and through these criteria the definition in (31) provides benchmarks for arguing this case. One interesting test case for such a discussion may be provided by the Nanti examples of self-quotation discussed in Section 2.3, which under a detailed semantic analysis along the lines of the definition proposed in (31) may or may not count as reported speech.

Finally, unlike (34), (35) does seem to illustrate a clear example of reported speech, but explaining why is not necessarily straightforward.

(35) "There is no tooth fairy," yeah right.

The quotation marks in (35) are an orthographic convention that can correspond to a variety of prosodic patterns, and there is no explicitly marked reported speaker or reported speech event. The definition in (31) can nevertheless easily identify (35) as reported speech: 'there is no tooth fairy' is a demonstrated element (an interpretation the quotation marks and prosody can help facilitate), it is necessarily an element the current speaker presents as having perceived in some perception event other than the current speech event, and the 'yeah right' following it is not random either: it specifies a modal attitude semantically prompted by the category of reported speech. As soon as each of these semiotic, evidential and modal features are interpretable, more explicit marking of the unit (e.g. by fully specifying M) can become redundant, but the essential categorial nature of this unit as reported speech is not determined by such marking choices.

The decision about whether the examples introduced in this section constitute reported speech *constructions* ultimately is a specific morpho-syntactic question about English, more particularly, about whether we can demonstrate within the confines of our particular syntactic model of choice that the structure of these examples sufficiently marks the meaning elements we have posited for reported speech in (31).[21] In this section our aims have been more general. We hope to have shown with the simple examples above that the elements in (31) allow for a structure-independent definition of reported speech that allows us to motivate why individual examples represent the category, or why not.

3.3 What does this definition solve?

In addition to categorising individual constructions as either reported speech or not, the definition presented in (31), we suggest, also allows us to account for all of the observations in Section 2.

21 We have explicitly refrained from adopting a specific syntactic paradigm in this paper, such as, for example, a functionalist layered model which, as one reviewer remarks, could be particularly helpful when describing diachronic developments in constructionalisation of reported speech. We agree, but hope that our present approach can contribute to a more cross-theoretical discussion, which we believe would be less readily facilitated by adopting a specific paradigm too early on.

In Section 2.1 we observed that the syntactic relation between the two elements in reported speech, which we labelled M and R, could not simply be reduced to a specific type of coordination or subordination. We suggest that this is one of the strongest arguments for defining reported speech at the level of a sentential construction, as our definition proposes, rather than at that of its individual constituent parts.

The main requirement of our defenestration analysis, i.e. the proposal that some elements of reported speech may be treated as structurally optional (see Section 2.2), is a stable semantic benchmark against which it can be judged which elements of a reported speech constructions need to be available. This is what our semantic definition of a reported speech construction in (31) aims to provide.

The inherent asymmetry between M and R (Section 2.3) follows from the role they have in the expression of the evidential meaning: M grounds the reported speech construction in the current speech situation, R represents a non-current speech event and its contents. In addition, R has a distinct semiotic status of being 'demonstrated'. The semiotic status of R may account for the observation that as a sub-clausal, clausal or even multi-clausal element R shows behaviour that is distinct from, e.g. regular subordinated or coordinated clauses and may contain pronouns and other construction types that are specific to R (Section 2.4).

The bi-perspectival meaning observed in Section 2.5 can be derived, on our definition, from the interaction between the inherent modal and evidential meanings of reported speech. The evidential meaning casts the speaker as a *perceiving participant* in the perception event (i.e. the reported speech event), while the modal meaning casts her as a participant evaluating the proposition (cf. Spronck 2015c). The contrast between these two epistemic roles, i.e. that of observer and that of 'evaluator', introduces a double perspective at the semantic core of reported speech. The interrelatedness between the semantic parameters in our definition and their relation to perspective interpretations is further demonstrated by the observations in Section 2.6: direct speech, e.g. presents R as a more distinct semiotic unit than, e.g. indirect speech, and is also associated with different modal and evidential values, resulting in the described attitudinal meanings.

In Section 2.7 we introduced several examples of apparent meaning change and constructionalisation in reported speech, resulting in interpretations as varied as 'to be about to do *p*', 'lest *p* occur' etc. As shown by the references introduced there, such examples are remarkably prolific. If a reported speech construction with semantic properties as in (31) were to constitute a source construction for grammaticalised/constructionalised examples as described in Section 2.7, we would expect to be able to relate their meaning to the semantic

components in (31). More specifically, we would predict that as these secondary functions of (former) reported speech constructions emerge, we would expect to see both further conventionalisation of form (i.e. more structural specialisation of constructions with a secondary meaning) as well as semantic bleaching or strengthening of the meaning components that are initially present in a reported speech construction. Strengthening of the 'demonstrated' meaning could result in, e.g. a new construction marking sentential units or prominence (Güldemann 2008; Jarque & Pascual 2016). Strengthening of the evidential meaning may result in a derived construction able to express non-speech related perception events, such as, e.g. inferential constructions. Strengthening of the modal meaning may result in attitudinal or epistemic constructions, like the Usan 'lest' construction in (29). Subtle combinations and recalibrations of these meaning components may result in yet other interpretations.[22] We believe that these hypotheses are compatible with the observed examples found in the literature referenced in Section 2.7, but acknowledge that this topic requires further diachronic and typological investigation.

Finally, (31) provides a conventional, semantic definition of a reported speech construction, and thereby challenges views in which the meaning of reported speech is emergent or primarily pragmatic. Our proposed semantic structure may even explain some of the extra-linguistic, multimodal realisations of reported speech referred to in Section 2.8, since both the demonstrated status and the indexicality inherent in the semantics of reported speech present semantic motivations for iconic and indexical gestures and changes in voice quality, postural mimicking, and other types of vocal and bodily imitation.

4 Conclusion

Our first aim in this paper has been to demonstrate that reported speech is a coherent, cross-linguistically regular phenomenon that displays features that cannot be derived from any other construction type. We believe that the observations reported here strongly suggest that reported speech constitutes a syntactic class in its own right, and in proposing a way to analyse the structures involved we hope to have contributed to further exploration of reported speech in a constructional typology.

22 For suggestions about how temporal and 'aspectual' meanings may arise out of modal and evidential meanings in reported speech, see Spronck (2016).

At a broader level, we also hope to have initiated a debate between very diverse schools of thought on linguistic meaning and language structure. Although we believe that the combination of arguments and analyses of reported speech laid out here have not been previously presented in a typological context, we do not wish to claim that most, or even any, of these are unique in the literature. Reported speech has been widely studied, within a range of traditions, and each of these traditions has had its own focus. For example, the semiotic, evidential and modal aspects of our definition of reported speech have all been suggested by various authors, but, as far as we are aware, have not been previously brought together. The questions of how to relate the meanings discussed to language structure go beyond the topic of reported speech in a strict sense and present challenges for interpreting, for example, perspectival meanings more generally. If our paper has been able to bring together aspects of these diverse traditions in a way that allows discussion of syntactic typology across boundary lines of individual theories, we will have succeeded in our second, and perhaps even more fundamental goal.

Funding: This project has received funding from the Faculty of Arts, University of Helsinki and the Helsinki University Humanities Programme (to Spronck) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 758232) (to Nikitina).

References

Aaron, Uche E. 1992. Reported speech in Obolo narrative discourse. In Shin Ja, J. Hwang & Willam R. Merrifield (eds.), *Language in context: Essays for Robert E. Longacre*, 227–240. Dallas: Summer Institute of Linguistics.

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2008. Semi-direct speech: Manambu and beyond. *Language Sciences* 30. 383–422.

Andrewes, M. 1951. The function of tense variation in the subjunctive mood of Oratio Obliqua. *The Classical Review* 1(3/4). 142–146.

Bally, Charles. 1912. Le style indirect libre en français moderne I & II. *Germanisch-Romanische Monatschrift* 4(549–556). 598–606.

Banfield, Ann. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan Paul.

Besnier, Niko. 1993. Reported speech and affect on Nukulaelae Atoll. In Jane H. Hill & Judith T. Irvine (eds.), *Responsibility and evidence in oral discourse*, 161–181. Cambridge: Cambridge University Press.

Blackwell, Natalia L., Marcus Perlman & Jean E. Fox Tree. 2015. Quotation as a multimodal construction. *Journal of Pragmatics* 81. 1–7. DOI: http://dx.doi.org/10.1016/j.pragma.2015.03.004. http://www.sciencedirect.com/science/article/pii/S0378216615000855.

Bogomolova, Natalia & Dima Ganenkov. 2010. Reported speech in Tabasaran: Person agreement patterns. Paper presented at Workshop on Reported speech in East Caucasian languages Lyon, 22 September 2010.

Bugaeva, Anna. 2008. Reported Discourse and Logophoricity in Southern Hokkaido Dialects of Ainu. *Gengo Kenkyu* 133. 31–75.

Burgess, Don. 1984. Western Tarahumara. In Ronald W. Langacker (ed.), *Studies in Uto-Aztecan Grammar 4*, 1–150. Dallas: Summer Institute of Linguistics.

Chappel, Hilary. 2012. Say in Sinitic: From Verba Dicendi to attitudinal discourse markers. In Johan van der Auwera & Jan Nuyts (eds.), *Grammaticalization and (inter-)subjectification*, 81–110. Wetteren: Universa Press.

Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123(3). 324–347. DOI: 10.1037/rev0000026.

Clark, Herbert H. & Richard J. Gerrig. 1990. Quotations as demonstrations. *Language* 66. 764–805.

Clements, George. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10. 141–177.

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Cornillie, Bert. 2009. Evidentiality and epistemic modality. On the close relationship between two different categories. *Functions of Language* 16(1). 9–43.

Coulmas, Florian (ed.). 1986a. *Direct and indirect speech*. Berlin: Mouton de Gruyter.

Coulmas, Florian. 1986b. Reported speech: Some general issues. In Florian Coulmas (ed.), *Direct and indirect speech*, 1–28. Berlin: Mouton de Gruyter.

Creissels, Denis. 2010. Person deixis in Northern Akhvakh reported speech.

Curnow, Timothy Jowan. 2002. Three types of verbal logophoricity in African languages. *Studies in African Linguistics* 31. 1 & 2.

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. http://dx.doi.org/10.1146/annurev-linguist-030514-125220.

Dancygier, Barbara & Eve Sweetser (eds.). 2012. *Viewpoint in language: A multimodal perspective*. Cambridge: Cambridge University Press.

Daniel, Michael. 2015. Logophoric reference in Archi. *Journal of Pragmatics* 88. 202–219.

Davidse K, L. Brems, P. Willemse, E. Doyen, J. Kiermeer, E. Thoelen, E. Miola (eds.). 2013. A comparative study of the grammaticalized uses of English "sort (of)" and French "genre (de)" in teenage forum data Languages Go Web. Standard and non-standard languages on the Internet, 41–66. Alessandria: Edizioni dell' Orso.

Davidson, Donald. 1979. *Quotation Theory and Decision* 11. 27–40.

Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38. 477–520. DOI: 10.1007/s10988-015-9180-1.

Deal, Amy Rose. 2017. *Shifty asymmetries: Universals and variation in shifty indexicality*.

De Brabanter, Philippe. 2017. Why quotation is not a semantic phenomenon, and why it calls for a pragmatic theory. In Ilse Depraetere & Raf Salkie (eds.), *Semantics and Pragmatics: Drawing a Line*, 227–254. Dordrecht etc.: Springer.

Deibler, Ellis. 1976. *Semantic relationships of Gahuku verbs* (SILPublications), vol. 48. Norman, OK: SIL of the University of Oklahoma.

Demirok, Ömer & Balkiz Öztürk. 2015. The Logophoric Complementizer in Laz. *Dilbilim Arastrmalari Dergisi* 2. 45–69.
Dingemanse, Mark & Akita Kimi. 2016. An inverse relation between expressiveness and grammatical integration: on the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics*, 1–32.
Dol, Philomena. 2007. *A Grammar of Maybrat: A language of the Bird's Head Peninsula, Papua Province, Indonesia*. Canberra: Pacific Linguistics.
Ershova, Ksenia. 2012. Reported speech and reportative grammaticalization in Besleney Kabardian. *Proceedings of the second Central European conference in linguistics for postgraduate students*, 71–87.
Etxepare, Ricardo. 2008. On quotative constructions in Iberian Spanish. In Ritva Laury (ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions*, 35–77. Amsterdam/Philadelphia: John Benjamins.
Evans, Nicholas. 2006. View with a view: Towards a typology of multiple perspective constructions. In Rebecca T. Cover & Yuni Kim (eds.), *Proceedings of the thirty-first annual meeting of the Berkeley Linguistics Society*, 93–120. Berkeley: Berkeley Linguistics Society.
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: Oxford University Press.
Frajzyngier, Zygmunt. 1991. The de dicto domain in language. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, 219–251. Amsterdam/Philadelphia: John Benjamins.
Ganenkov, Dima. 2010. Reported speech in Dargwa: An overview. Paper presented at Workshop on Reported speech in East Caucasian languages Lyon, 22 September 2010.
Gawne, Lauren. 2013. Lamjung Yolmo copulas in use: Evidentiality, reported speech and questions. Department of Linguistics and Applied Linguistics, School of Languages and Linguistics, The University of Melbourne.
Gawne, Lauren. 2015. The reported speech evidential particle in Lamjung Yolmo. *Linguistics of the Tibeto-Burman Area* 38(2). 292–318.
Gentens, Caroline. 2016. The factive-reported distinction in English. The semantics, construal and interpersonal status of finite complement clauses, University of Leuven.
Gildersleeve, Basil L. 1906. Notes on the Evolution of Oratio Obliqua. *The American Journal of Philology* 27(2). 200–208.
Ginzburg, Jonathan & Robin Cooper. 2014. Quotation via Dialogical Interaction. *Journal of Logic, Language and Information* 23(3). 287–311. DOI: 10.1007/s10849-014-9200-5.
Greenough, J. B., A. A. Howard, G. L. Kittredge & Benj. L. D'Ooge. (eds.). 1903. *Allen and Greenough's New Latin grammar for schools and colleges: Founded on comparative grammar*. Boston/London: Ginn.
Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.
Güldemann, Tom. 2012. Thetic speaker-instantiating quotative indexes as a cross-linguistic type. In Isabelle Buchstaller & Ingrid van Alphen (eds.), *Quotatives: Cross-linguistic and cross-disciplinary perspectives*, 117–142. Amsterdam/Philadelphia: John Benjamins.
Güldemann, Tom & Manfred von Roncador (eds.). 2002. *Reported discourse: A meeting ground for different linguistic domains*. Amsterdam/Philadelphia: John Benjamins.
Hagège, Claude. 1974. Les pronoms logophoriques. *Bulletin De La Societe De Linguistique De Paris* 69. 287–310.

Haller, Felix & Chungda Haller. 2007. Einführung in das moderne Zentraltibetische. Auf Basis des Dialektes von Shigatse / westliches Zentraltibet (Tsang).

Haßler, Gerda. 2002. Evidentiality and reported speech in Romance languages. In Tom Güldemann & Manfred von Roncador (eds.), *Reported discourse, A meeting ground for different linguistic domains*, 143–172. Amsterdam/Philadelphia: John Benjamins.

Haßler, Gerda, Gabriele Diewald & Elena Smirnova. (eds.). 2010. *Epistemic modality and evidentiality and their determination on a deictic basis: the case of Romance languages Linguistic Realization of Evidentiality in European Languages*, 223–248. Berlin/New York: De Gruyter Mouton.

Hedinger, Robert. 1984. Reported speech in Akɔɔse. *Journal of West African Languages* XIV. 81–102.

Heine, Bernd & Kuteva Tania. 2002. *World lexicon of grammaticalization Cambridge etc.*: Cambridge University Press.

Hellwig, Birgit. 2006. *Complement clause type and complementation strategies in Goemai. Complementation*, 204–223. Oxford: Oxford University Press.

Hengeveld, Kees & Marize Mattos Dall'Aglio Hattnher. 2015. Four types of evidentiality in the native languages of Brazil. *Linguistics* 53(3). 479–524.

Hercus, Luise A. 1994. *A grammar of the Arabana-Wangkangurru language, Lake Eyre basin, South Australia*. Canberra: Pacific Linguistics.

Hill, Harriet. 1995. Pronouns and reported speech in Adioukrou. *Journal of West African Languages* XXV. 87–106.

Hill, Jane H. 1995. The voices of Don Gabriel: Responsability and self in a modern Mexicano narrative. In Dennis Tedlock & Brice Mannheim (eds.), *The Dialogic Emergence of Culture*, 97–147. Urbana/Chicago: University of Illinois Press.

Jakobson, Roman. 1957. *Shifters, verbal categories and the Russian verb*. Department of Slavic Languages and Literatures. Cambridge, MA: Harvard University.

Janssen, Theo & Wim van der Wurff (eds.). 1996. *Reported speech: Forms and functions of the verb*. Amsterdam/Philadelphia: John Benjamins.

Jarque, Maria Josep & Esther Pascual. 2016. Mixed viewpoints in factual and fictive discourse in Catalan Sign Language narratives. In Barbara Dancygier, Lu Wei-Lun & Arie Verhagen (eds.), *Viewpoint and the fabric of meaning: Form and use of viewpoint tools across languages and modalities*, 259–280. Berlin/Boston: De Gruyter Mouton.

Kaplan, David, Joseph Almog, John Perry & Howard Wettstein (eds.). 1989. Demonstratives Themes from Kaplan, 481–563. Oxford etc.: Oxford University Press.

Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in linguistics*, 143–173. The Hague: Mouton.

Kolyaseva, Alena F. 2018. The 'new' Russian quotative tipa: Pragmatic scope and functions. *Journal of Pragmatics* 128. 82–97.

Lampert, Martina. 2013. Say, be like, quote (unquote), and the air-quotes: Interactive quotatives and their multimodal implications. *English Today* 29(04): 45–56. DOI: 10.1017/S026607841300045X. http://journals.cambridge.org/article_S026607841300045X.

Larson, Mildred Lucille. 1978. *The functions of reported speech in discourse*. Dallas: Summer Institute of Linguistics.

LaTerza, Chris, Morgan Rood, Dustin Chacón, Ruth Kramer & Jen Johnson (ed.). 2015. *New puzzles for shifting indexicals: An Amharic case study selected proceedings of 38 Stef Spronck and Tatiana Nikitina the 44th annual conference on African linguistics*, 158–164. Somerville, MA: Cascadilla Proceedings Project.

Lips, Marguerite. 1926. *Le style indirect libre*. Paris: Payot.
Lucy, John A. 1993. Reflexive language and the human disciplines. In John A. Lucy (ed.), *Reflexive language: Reported speech and metapragmatics*, 9–32. Cambridge etc.: Cambridge University Press.
Maier, Emar. 2017. The pragmatics of attraction: Explaining unquotation in direct and free indirect discourse. In P. Saka & M. Johnson eds., *The Semantics and Pragmatics of Quotation*, 259–280. Cham: Springer.
Maisak, Timur & Solmaz Merdanova. 2010. Reported speech and "semi-directness" in Agul.
Malibert, Il-Il & Martine Vanhove. 2015. Quotative constructions and prosody in some Afroasiatic languages: Towards a typology. In Amina Mettouchi, Martine Vanhove & Dominique Caubet (eds.), *Corpus-based studies of lesser-described languages: The CorpAfroAs corpus of spoken AfroAsiatic*, 117–169. Amsterdam/Philadelphia: John Benjamins.
Matic, Dejan & Brigitte Pakendorf. 2013. Non-canonical SAY in Siberia. *Studies in Language* 37(2). 356–412.
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.
McGregor, William B. 2011. *The Nyulnyul language of Dampier Land, Western Australia*. vol. 1 and 2. Canberra: Pacific Linguistics.
McGregor, William B. 2013. Optionality in grammar and language use. *Linguistics* 51(6). 1147–1204.
McGregor, William B. 2014. The "say, do" verb in Nyulnyul, Warrwa, and other Nyulnylan languages is monosemic. In Klaus Robering (ed.), *Events, arguments and aspects: Topics in the semantics of verbs*, 301–327. Amsterdam/Philadelphia: John Benjamins.
Michael, Lev. 2010. Rethinking quotatives, reported speech, and Uutterance responsibility: Implications of Nanti self-quotation.
Michael, Lev. 2014. Nanti self-quotation: Implications for the pragmatics of reported speech and evidentiality. In Janis Nuckolls & Lev Michael (eds.), *Evidentiality in interaction*, 155–191. Amsterdam/Philadelphia: John Benjamins.
Munro, Robert, Rainer Ludwig, Uli Sauerland & David W. Fleck. 2012. Reported speech in matses: Perspective persistence and evidential narratives. *International Journal of American Linguistics* 78(1). 41–75.
Nau, Nicole. 2006. Out of Africa: Logophoric pronouns and reported discourse in finnish and high Latvian dialects. *Acta Linguistica Lithuanica* LV. 55–87.
Nikitina, Tatiana. 2012a. Personal deixis and reported discourse: Towards a typology of person alignment. *Linguistic Typology* 16(2). 233–263.
Nikitina, Tatiana. 2012b. Logophoric discourse and first person reporting in Wan (West Africa). *Anthropological Linguistics* 54(3). 280–301.
Nikitina, Tatiana. 2012c. Russian verboids: A case study in expressive vocabulary. *Linguistics* 50. 165–189.
Nikitina, Tatiana. 2018. When linguists and speakers do not agree: The endangered grammar of verbal art in West Africa. *Journal of Linguistic Anthropology* 28(2). 197–220.
Novelli, B. 1985. *A grammar of the Karimojong language*. Berlin: Reimer.
Pascual, Esther. 2014. *Fictive interaction: The conversation frame in thought, language, and discourse*. Amsterdam/Philadelphia: John Benjamins.

Plank, Frans. 1986. Über den Personenwechsel und den anderer deiktischer Kategorien in der wiedergegebenen Rede. *Zeitschrift für germanistische Linguistik* 14. 284–308.
Postgate, J. P. 1905. Repraesentatio temporum in the Oratio Obliqua of Caesar. *The Classical Review* 19. 441–446.
Recanati, François. 2001. Open quotation. *Mind* 110(439). 637–687.
Reesink, Gerard P. 1993. Inner speech' in Papuan languages. *Language and Linguistics in Melanesia* 24. 217–225. http://www.sil.org/resources/archives/22744.
Reid, Aileen A. 1979. Dynamics of reported speech in Totonac. In Linda K. Jones (ed.), *Discourse studies in Mesoamerican languages*, 293–328. Dallas: Summer Institute of Linguistics.
Roeck, Marijke De. 1994. A functional typology of speech reports. In E. Engberg-Pedersen, L. Jakobson & S. Rasmussen (eds.), *Function and expression in universal grammar*, 331–351. Berlin/New York: De Gruyter Mouton.
Rumsey, Alan. 1990. Wording, meaning and linguistic ideology. *American Anthropologist* 92(2). 346–361.
Rumsey, Alan. 2010. 'Optional' ergativity and the framing of reported speech. *Lingua* 120. 1652–1676.
Saka, Paul. 1998. Quotation and the use-mention distinction. *Mind* 107(425). 113–135. DOI: 10.1093/mind/107.425.113.
Salmon, E. T. 1931. A note on subordinate clauses In Oratio Obliqua. *The Classical Review* 45. 173.
Saxena, Anju. 1988. On syntactic convergence: The case of the verb "say" in Tibeto-Burman. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 375–388.
Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26. 29–120.
Sergeeva, E. N. Kazanskij, N. N. (ed.). 2010. *Tipa kak èvidential'nyj marker v russkoj razgovornoj reči. Acta Linguistica Petropolitana: Trudy instituta lingvističeskix issledovanij* IV. 149–153.
Shklovsky, Kirill & Yasutada Sudo. 2014. The Syntax of Monsters. *Linguistic Inquiry* 45. 381–40.
Si, Aung & Stef Spronck. Forthcoming. Solega defenestration: The grammar of underspecified perspective shift in Solega (Dravidian). *Pragmatics*.
Silverstein, Michael. 1976. Hierarchy of features and ergativity. In M.W. Dixon Robert (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.
Skjon, E. Lee. 2001. Point of view, reported speech, and obviation in two Fox texts. In J. D. Nichols (ed.), *Actes du Trente-Deuxième Congrés des Algonquinistes*, 493–515. Winnipeg: University of Manitoba.
Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114. 255–276.
Spronck, Stef. 2009. How different is reportativity from reported speech? In Andrea Ender, Marc Matter & Fabienne Tissot (eds.), *Proceedings der 39. Studentische Tagung Sprachwissenschaft (STuTS) in Bern* (Arbeitspapiere), vol. 45, 3–19. Bern: Department of Linguistics.
Spronck, Stef. 2012. Minds divided, speaker attitudes in quotatives. In Isabelle Buchstaller & Ingrid Van Alphen (eds.), *Quotatives: Cross-linguistic and cross-disciplinary perspectives*, 71–116. Amsterdam/Philadelphia: John Benjamins.

Spronck, Stef. 2015a. Reported speech in Ungarinyin: Grammar and social cognition in a language of the Kimberley region, Western Australia. The Australian National University. http://hdl.handle.net/1885/733712596.

Spronck, Stef. 2015b. Refracting views: How to construct complex perspective in reported speech and thought in Ungarinyin. *STUF – Language Typology and Universals* 68(2). 165–185.

Spronck, Stef. 2015c. Stance as participant structure: A Jakobsonian approach to the pragmatics and semantics of evidentiality. *Belgian Journal of Linguistics* 29. 193–216.

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. 2017. Defenestration: Deconstructing the frame-in relation in Ungarinyin. *Journal of Pragmatics* 114. 104–133. DOI: https://doi.org/10.1016/j.pragma.2017.03.016. http://www.sciencedirect.com/science/article/pii/S0378216616304398.

Stec, Kashmiri, Mike Huiskes & Gisela Redeker. 2015. Multimodal analysis of quotation in oral narratives. *Open Linguistics* 1. 531–554.

Stephenson, Tamina. 2007. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30. 487–525.

Tannen, Deborah. 2007. *Talking voices: repetition, dialogue and imagery in conversational discourse*, 2nd ed. Cambridge: Cambridge University Press.

van der Voort, Hein. 2002. The quotative construction in Kwaza and its (de-) grammaticalisation. In Mily Crevels, Simon van de Kerke, Sérgio Meira & Hein van der Voort (eds.), *Current studies on South American languages* [Indigenous Languages of Latin America, 3], 307–328. Leiden: Research School of Asian, African, and Amerindian Studies (CNWS).

Vandelanotte, Lieven. 2004. Deixis and grounding in speech and thought representation. *Journal of Pragmatics* 36(3): 489–520. DOI: http://dx.doi.org/10.1016/j.pragma.2003.10.003. http://www.sciencedirect.com/science/article/pii/S0378216603001437.

Vandelanotte, Lieven. 2008. Dependency, framing, scope? The syntagmatic structure of sentences of speech or thought representation. *Word* 59(1). 55–82.

Vandelanotte, Lieven. 2009. *Speech and thought representation in English: A cognitive-functional approach*. Berlin/New York: De Gruyter Mouton.

Verstraete, Jean-Christophe. 2011. The functions of represented speech and thought in Umpithamu narratives. *Australian Journal of Linguistics* 31(4). 491–517.

Vološinov, Valentin N. 1930. *Marksizm i filosofija jazyka*, 2nd ed. Leningrad: Priboj.

Vološinov, Valentin N. 1973. Ladislav Matejka & I.R. Titunik (eds.), *Marxism and the philosophy of language*. New York/London: Seminar Press.

Widmer, Manuel & Marius Zemp. 2017. The epistemization of person markers in reported speech. *Studies in Language* 41(1). 33–75. DOI: 10.1075/sl.41.1.02wid.

Wiemer, Björn, Diewald, Gabriele & Smirnova, Elena. (eds.). 2010. *Hearsay in European languages: toward an integrative account of grammatical and lexical marking Linguistic Realization of Evidentiality in European Languages*, 59–129. Berlin/New York: De Gruyter Mouton.