



In a Manner of Speaking: How Reported Speech May Have Shaped Grammar

Stef Spronck* and Daniela Casartelli

University of Helsinki, Helsinki, Finland

We present a first, broad-scale typology of extended reported speech, examples of lexicalised or grammaticalised reported speech constructions without a regular quotation meaning. These typically include meanings that are conceptually close to reported speech, such as THINK OR WANT, but also interpretations that do not appear to have an obvious conceptual relation with talking, such as CAUSE OR BEGIN TO. Reported speech may therefore reflect both concepts of communication and inner worlds, and meanings reminiscent of ‘core grammar’, such as evidentiality, modality, aspect (relational) tense and clause linking. We contextualise our findings in the literature on fictive interaction and perspective and suggest that extended reported speech may lend insight into a fundamental aspect of grammar: the evolution of verbal categories. Based on the striking similarity between the meanings of extended reported speech and grammatical categories, we hypothesise that the phenomenon represents a plausible linguistic context in which grammar evolved.

OPEN ACCESS

Edited by:

Sonja Zeman,
Ludwig Maximilians Universität
München, Germany

Reviewed by:

Sergeiy Sandler,
Independent researcher, Beer Sheva,
Israel

Henrik Bergqvist,
Stockholm University, Sweden

*Correspondence:

Stef Spronck
stef.spronck@helsinki.fi

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 31 October 2020

Accepted: 12 July 2021

Published: 07 September 2021

Citation:

Spronck S and Casartelli D (2021) In a
Manner of Speaking: How Reported
Speech May Have Shaped Grammar.
Front. Commun. 6:624486.
doi: 10.3389/fcomm.2021.624486

Keywords: reported speech, quotation, perspective, TAME, fictive interaction, evolution of language, linguistic typology

1 INTRODUCTION: FICTIVE INTERACTION, REPORTED SPEECH AND GRAMMAR

The act of speaking is so fundamental to the human experience and our perception of other people that we routinely cast our interaction with the world as a dialogue (Bakhtin, 1981). We may represent intentions of others, even those attributed to non-humans, as ‘speaking to us’, as in (1a) or (1b). In these sentences a perception or, e.g., the expression of the face of an animal is described as a speech event.

- (1) a. One look said it all.
b. Its wrinkled snout told me: “Don’t come near!”

In a series of seminal accounts (Pascual, 2002; Pascual, 2007 and, especially, Pascual, 2014) characterises expressions as in (1) as ‘fictive interaction’, defined as the use of ‘conversation as a frame to structure mental, discursive, and linguistic processes’ (Pascual, 2014, 9). While this analysis is explicitly embedded in a cognitive linguistic account that sees conversation as a cognitive *Gestalt* that humans may use in order to make sense of the world, its empirical foundation is strong. Not only are examples of fictive interaction as in (1) common cross-linguistically (Pascual and Sandler, 2016a; McGregor, 2019), they affect a heterogeneous set of sentence types and linguistic structures (Pascual and Sandler, 2016b).

While in many Standard Average European languages fictive interaction appears to be a metaphor-driven type of creative language use (apart from possibly more idiomatic expressions like to ‘speak to oneself’), a much more conventionalised form of fictive interaction occurs in many languages around the world, cf. (2).

(2) WAN (Mande; Ivory Coast)

yī ē gé bā kó
water DEF say LOG boil

‘The water was about to boil.’ (lit., ‘The water said: let me boil!’) (Nikitina, 2020, 88)

Like in (1), (2) presents a ‘speech event’ that does not involve actual communication. The literal translation of (2) clearly demonstrates this: the water incites its own boiling, as if it was speaking. Crucially, however, this is not what (2) *means*. As the idiomatic gloss in (2) illustrates, the interpretation of this sentence is an inceptive meaning, i.e. ‘to be about to do *p*’. Rather than *saying* “let me boil”, as the lexicogrammatical elements in (2) suggest, the only plausible contextual meaning of the sentence is ‘The water was about to boil’.

Whereas it is intuitively obvious to the reader that the interpretation of the ‘fictive’ example of direct speech in (1b) is metaphorical and is interpreted through a process of inference, the structural, semantic and pragmatic status of examples like (2) are much less clear. This is problematic, because this is not an isolated or anecdotal example. Examples that, judging by their lexico-grammatical components would seem to carry a meaning of direct speech (x said: “*p*”), may receive widely varying interpretations in the languages of the world. These interpretations include meanings that seem to have little in common with speech events, or the perspective-shifting function associated with reported speech.

In a first cross-linguistic typology of the phenomenon, specifically focusing on direct speech-like structures as in (2), Pascual (2014, 90) distinguishes meanings as varied as (1) mental states, (2) emotional and attitudinal states, (3) desires, (4) intentions, (5) attempts, (6) states of affairs¹, (7) causation, (8) reason, (9) purpose and (10) future tense. We may group these meanings into the four classes in (3).

(3) THOUGHT: mental states, desires, intentions, attempts;
ATTITUDE: emotional states;
CAUSE: causation, reason, purpose;
TIME: states of affairs, future tense.

Despite the extensive list of meanings Pascual’s pioneering study uncovers, it raises several important questions. First of all: what is the status of the meanings in (3)? In order to answer this question we will need to understand, firstly, if the list in (3) is exhaustive with respect to the range of meanings attested in similar examples so far, secondly, if these meanings are random or show recurrent patterns in the languages of the world and, thirdly, what mechanisms give rise to the meanings as in (3)? These are questions the current article aims to address.

¹Pascual (2014) uses this label to describe interpretations relating to the internal organisation of an event (particularly inchoative meanings; see Section 2.2.2). This class roughly corresponds to the function we will describe as ‘aspect’ below.

This objective immediately faces a methodological challenge: both in (1b) and (2), the lexico-grammatical make-up of the examples suggests a ‘literal’ direct speech interpretation, but the actual meaning in context varies, as well as the way in which this meaning arises pragmatically. As we will see below, the semantically based notion of ‘direct speech’ does not neatly apply to all relevant instances. This suggests that in order to examine the variation of the meanings involved, we need to start with a definition of a class of relevant examples based on their lexico-grammatical properties. For the sake of cross-linguistic comparison, this set of lexico-grammatical properties cannot be too restrictive, since it needs to be applicable to languages of distinct structural types. We cannot make *a priori* assumptions about the language-specific variation that might exist between these structures. On the other hand, it needs to capture a class of phenomena that are cross-linguistically comparable and can be identified based on the definition, so it cannot be too inclusive either (Haspelmath, 2010). We will return to the wider context of fictive interaction at the end of this article, but in order to maximally avoid the presumption of metaphoricity implicit in this label we will refer to the typological examples examined in this article as ‘extended reported speech’. The identification of relevant reported speech examples will be based on the definition in (4a). Extended reported speech will be defined as in (4b).

- (4) a. For the purposes of this study, a reported speech construction is a clausal or bi-clausal linguistic structure consisting of a Matrix unit (M) that has scope over a Report unit (R) that, at least under one of its interpretations, can be understood as a representation of an utterance. M minimally consists of or contains an element that can be translated as ‘say’.
b. For the purposes of this study, extended reported speech is a reported speech construction that is *not* interpreted as representing a speech act.

We begin in **Section 2** with an extensive illustration of the definitions, or ‘comparative concepts’ (Haspelmath, 2010), in (4a) and (4b), showing how they can be applied across languages and what type of examples they unveil. These illustrations may also clarify some of the specific formulations in the definitions above, so we will address further motivations for the comparative concepts in (4a) and (4b) in **Section 2**. The section begins with a brief contextualisation of the typological and descriptive literature on which we will draw (**Section 2.1**), before exemplifying some of the main attested types of extended reported speech in **Section 2.2**. These observations both support and expand the initial classification in (3), and **Section 2.3** presents an updated list of extended reported speech meanings.

As we will show, two meanings that seem particularly well documented so far in extended reported speech constructions are those with an intention reading (i.e. a ‘want’ meaning) and those with a complementiser meaning. These are two types we will explore in more detail in **Section 3**, based on a cross-linguistic sample of 100 genetically diverse languages. These are the first results of a sample study aiming to develop a broad-scale typology of extended reported speech. The sample and methodology of the study are introduced in **Section 3.1** and **Section 3.2** presents and illustrates the results. The observations and initial analyses from **Sections 2** and **3** are summarised and integrated in **Section 3.3**.

With this empirical foundation in place, in **Section 4** we suggest some implications of our observations for the understanding of the semantics of reported speech, perspective constructions, particularly the subtypes of reported speech and, speculatively, the evolution of grammar. In **Section 4.1** we relate the extensions of reported speech to the semantics of reported speech constructions and identify three pathways towards extended reported speech, specifically, recasting, rescaling and semantic bleaching. The implications of extended reported speech for our understanding of the diachrony of perspective expressions in language and theories of classifications of direct and indirect speech are explored in **Section 4.2**. We outline our main motivation behind this project in **Section 4.3**, where we claim that the observations about extended reported speech demonstrate so many similarities with the meaning of common grammatical categories that the phenomenon holds fundamental implications for the evolution of grammar.

Finally, **Section 5** presents a brief conclusion.

2 STUDIES ON EXTENDED REPORTED SPEECH: A SURVEY

2.1 Background

Extended reported speech has been relatively well documented in the descriptive and typological literature. One of the earliest in-depth studies of the phenomenon appears to be Larson (1978), who discusses reported speech in the South-American language Aguaruna, demonstrating that it can be used to express meanings far beyond speech representation, cf. (5).

- (5) AGUARUNA (Chicham; Perú)
- a. *Wetatjai, "Pujusti," tau san.*
I-will-go that-he-stay one-who-said I-being
Meaning in context: 'I will go because I want him to stay.' (Larson, 1978, 93)
Mock English gloss: '[I will go] "that he stays," I am the one who says'
- b. *Shutuktut tawa nunak ayujut taku tawai.*
shutuktut it-says that-obj-topic give-to-eat saying it:says
Meaning in context: 'Shutuktut' means to give someone something to eat.' (Larson, 1978, 113)
Mock English gloss: "Shutuktut" it says that "give to eat" saying it says

Following the definitions in (4), the examples in (5) illustrate extended reported speech since they both contain Report and Matrix units that, as the glosses illustrate, can be interpreted as representing reported utterances and clauses of saying, respectively (cf. 4a). Yet, as the idiomatic glosses (i.e., the third line of the examples) illustrate, the contextual interpretation of these examples does not involve a speech event (cf. 4b). This is how we will apply the definitions throughout this study: the comparative concept of a reported speech construction (4a) is evaluated against the morphemic gloss (i.e., the second lines of the examples), that of extended reported speech (4b) against the idiomatic gloss (i.e., the third lines of the examples).

In order to increase readability, we will also add a fourth line to each example, as in (5). This line is a mock English gloss that represents what the example could be expected to mean based on its lexico-grammatical content, i.e. it is a prose interpretation of

the morphemic glosses². Crucially, however, the Mock English gloss should not be taken to indicate the actual meaning of the full example; it is a presentational device in order to make the morphemic gloss more accessible³. In order to highlight this interpretative status, the fourth line also appears in a different font, below the translation given in the source. Elements placed between curly brackets in the Mock English glosses (as in 5a) are not part of the extended reported speech construction.

Apart from in Aguaruna, extended reported speech has been attested in languages across South America (van der Voort, 2002; Everett, 2008; Birchall, 2018). Several studies have described it as a regional phenomenon, occurring in languages in the Tibetan area (Saxena, 1988), in Africa (Güldemann, 2008), among Sinitic languages (Chappell, 2012) and across Siberia (Matić and Pakendorf, 2013). Furthermore, numerous studies of extended reported speech in Australia (Rumsey, 1990; McGregor, 2014, cf.), Austronesia and Papunesia (Deibler, 1971; Reesink, 1993; Klamer, 2000, cf.) and Central Asia Baranova (cf. 2015) have established it as a common phenomenon in languages of these areas as well. **Figure 1** shows the location the languages cited in this section, illustrating that descriptions of extended reported speech are not restricted to any particular geographical area or language family⁴.

Given how widespread the phenomenon of extended reported speech appears to be across the languages of the world, it is not surprising that it can carry many diverse meanings. What calls for an explanation, however, is the observation that these meanings, while wide-ranging, seem far from random. **Figure 2** summarises the most frequently occurring meanings described in the studies on extended reported speech that we will survey in this section.

The labels in **Figure 2**, which we will refer to as 'functions' or 'interpretations', represent a short, standardised summary of the meaning description given in each of the sources. The descriptions and classifications for the individual languages are listed in **Appendix Table A1**. Following this standardised list, (5a) carries a WANT function and (5b) a NAME function. Throughout this section we will introduce and illustrate each of the functions in **Figure 2**, as well as a few less commonly described ones.

²Since the segmentation and labelling in the morphemic gloss (i.e., the second line in all examples) reflect careful analytical choices on the part of the individual authors cited, most glosses in the examples introduced here have been preserved from the original source reference. Where these include abbreviations that do not follow the standard of the Leipzig Glossing Rules (Bickel et al., 2008) these are listed in the glossary at the end of this article. The only exception to this practice has been glosses that conflict with those in the Leipzig Glossing Rules, as in (13a), which uses 'S' for 'singular', whereas it indicates an intransitive subject in Bickel et al. (2008). This example also contains the gloss 'DEC' for 'declarative', which is minimally distinct from the standardised gloss 'DECL'. In such cases we have revised the glosses in accordance with the Leipzig Glossing Rules and have explicitly indicated this in the reference by adding 'gloss updated'.

³We will avoid the misleading term 'literal meaning' to refer to this line, since it suggests that the meaning indicated in the translation line (i.e., the third lines of the examples) is a metaphorical interpretation of the Mock English translation, which we do not assume to be the case for these examples.

⁴The larger dots on the map represent areal studies, which include multiple languages near the indicated location.



FIGURE 1 | Map of locations of languages cited in this section.

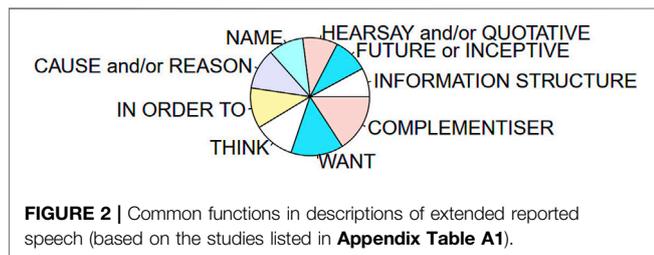


FIGURE 2 | Common functions in descriptions of extended reported speech (based on the studies listed in **Appendix Table A1**).

In order to provide an initial grouping into separate types, we will distinguish between examples of extended reported speech that, although conventionalised, are still clearly identifiable as reported speech constructions (**Section 2.2.1**) and those that show more signs of grammaticalisation (**Section 2.2.2**). Since we will focus on the functions of these examples and languages may show varying degrees of grammaticalisation in their forms of extended reported speech, these groupings are not entirely clear-cut, nor mutually exclusive, but they allow us to most clearly connect with existing descriptions in the literature.

2.2 Examples of Extended Reported Speech

2.2.1 Lexicalised and Conventionalised Examples

Perhaps the most common type of extended reported speech is examples with a ‘think’ interpretation. On a trivial level, this use of reported speech may seem familiar to Standard Average European

expressions like ‘I would say *p*’, which signals ‘I think *p*’, and, naturally, saying *p* implies thinking *p*. However, the extended reported speech version of this interpretation arises in languages in which the distinction between reported speech and reported thought is principally underspecified, cf. (6).

(6) KAVALAN (Austronesian; Taiwan)

“sessen=ti sunis unay,” zin-ku; bula-an-ku tu qudus-ku.
 AF.cold=PFV child that say-1SG.GEN give-LF-1SG.GEN OBL clothes-1SG.GEN
 ‘“That child was cold,” I thought; I gave my clothes to him.’ (Hsieh, 2012, 467)
 ME: “That child was cold,” I said. {I gave my clothes to it}

In (6), the verb that constitutes the Matrix unit, while glossed as ‘say’, could equally mean ‘say’ or ‘think’. Hsieh (2012), 467 writes about this example: ‘when no obvious addressee can be found in the clause, this may pose some difficulties in deciding whether the term in question denotes an act of speaking or an act of thinking. The correct interpretation depends heavily on pragmatic inferences’. In the languages for which reported thought has been described as a function of extended reported speech, the absence of an explicit reported addressee appears to be a common prompt for a thought interpretation (also cf. Spronck, 2015, 1–2). Particularly in languages in which the verb used in the Matrix unit does not indicate a strict lexical distinction between ‘say’, ‘think’ and, e.g., a generic action, as is the case in several Australian (Rumsey, 1990; McGregor, 2014) and South American languages (van der Voort, 2002), reported speech and reported thought are often virtually indistinguishable. This is also the case for many of the examples in the African languages Güldemann

(2008) describes under the label of ‘quotative indexes’. These are Matrix units, often consisting of a single morpheme, that typically (diachronically) derive from a lexeme meaning ‘say’, but that, synchronically, have a much broader meaning⁵.

As Hsieh (2012) suggests, the interpretation process involved in extended reported speech with a THINK function is often one based on inference, but there is a crucial difference with SAE examples like ‘to say to oneself’: the THINK examples often do not strictly codify the distinction between reported speech and reported thought: the example in (6) strictly expresses neither reported speech or thought; it can equally express both. No language has been reported to have a dedicated reported thought construction extending to a speech meaning. That is, for all languages, the speech interpretation appears to be the most common and versatile. However, the lexico-grammatical structure does not unambiguously specify this. Hence, the inferential interpretation *narrows* its meaning to THINK, rather than metaphorically extends it from a specific speech interpretation to a thought interpretation. The absence of a (clear) second referent indicating a person spoken to in such cases, suggests an undirected monologue, which leads to the interpretation that the subject referent of the Matrix unit is thinking the content of the Report unit, rather than saying it.

THINK-type extended reported speech appears common in the literature and some authors even assume that it underlies other subsequent meaning extensions illustrated below. For example, Reesink (1993) suggests that all extended reported speech could be seen as a form of ‘inner speech’, a term first coined by Vygotsky (1987), to reflect the idea that verbalised (but non-spoken) thought is like speaking in one’s mind. While this connection highlights the universal human cognitive principles behind the phenomenon, the metaphorical extension from SAY to THINK in, e.g., Standard Average European languages should not be confused with extended reported speech as intended here. For languages that do display the phenomenon as defined in (4) and exemplified in this section, THINK could be seen as the first stage crossing the Rubicon from ‘regular’ reported speech to extended reported speech.

A type slightly further removed from this stage is formed by the ‘intention’ interpretation of extended reported speech, which often can be translated with a lexeme meaning WANT. An example of this type is shown in (7) and in (5a) above.

(7) WARRWA (Nyulnyulan; Australia)

miliya ka-nda nga-n-di-ny-jarri marlu nga-la-nda-na
today IMIN.NOM.FUT-go IMIN.NOM-CM-say-IMP-SEQ not IMIN.NOM-IRR-go-PST

‘I wanted to go today, but didn’t’ (McGregor, 2014, 310)

ME: ‘I will go today,’ I said, {but I did not go}

⁵Both the observation that the main element in the Matrix unit can have a broader lexical meaning than ‘say’ and does not even need to be a fully inflecting verb, as is the case for many of Güldemann’s ‘quotative indexes’ motivate the inclusive formulation in our comparative concept in 1 that ‘M minimally consists of or contains an element that can be translated as ‘say’.

The way in which the Warrwa and Aguaruna strategies in (5a) and (7) are interpreted may rely on a similar inferential process as described for THINK: both examples are semantically underspecified. In the absence of an explicit reported addressee, like with THINK, (5a) and (7) suggest a monologic or internal process. Furthermore, in both instances the Report unit describes a future event with a first person subject, which seems appropriate for an intentional interpretation. Note again, however, that as with all examples of extended reported speech, the meanings of (5a) and (7) are those of the idiomatic glosses: even though we may be able to understand some of the compositional elements that give rise to the ‘want’ interpretation, these constructions are either the only, or a common way to express WANT complement constructions in the respective languages. (For similar observations about the grammatical status of ‘intentional’ reported speech constructions, see Rumsey (1990), Everett (2008) and Konnerth (2020), among others.) We return to this type in more detail in Section 3.2.

A final type of extended reported speech in which the apparent reported speaker is engaged in a mental rather than a speech activity is a broad class of attitudinal meanings that several authors discuss. Two relevant examples occur in (8).

(8) USAN (Trans–New Guinea; Papua New Guinea)

a. *me ya-g-ab eng yag umo-nd is-orei qamb ye-toat is-umirei*
not me-see-SS the river down-at:the descend-3s:FP say:SS me-follow:SS descend-3s:FP
‘...and when they didn’t see me they followed me thinking I had gone down to the river’
(Reesink, 1993, 222)

ME: {they did not see me down at the river and descended,} they said they followed me

b. *Mi qei-qei mani umer-iner qamb gitab ig-oun*
thing some-RED yam wilt-3s:UF say:SS abstain:SS be-1p:PRS
‘We abstain from various things lest the yam wilts’ (Reesink, 1993, 222, gloss updated)

ME: “‘the yam might wilt’” we say, {and abstain}

Example (8a) could be interpreted as an example of reported thought, but Reesink (1993) suggests that while the sentence attributes the thought that the current speaker had descended towards the river to the subject of *qamb* ‘they say’ in (8a), the example primarily conveys that this thought was mistaken, not that it was held (or uttered). In this survey, we will not explore this type beyond these observations, but attitudinal meanings are more commonly described in the literature on extended reported speech and the irrealis interpretation reported for Sinitic (Chappell, 2012) may be related to this as well.

The attitudinal meaning is perhaps even more explicit in (8b), since the idiomatic translation does not even include a cognitive or utterance verb. This meaning seems related to the interpretation of ‘warning’, which van der Voort (2002) lists for Kwaza and ‘deontic modality’, which Güldemann (2008) describes for his African sample.

We will return to the more general principles behind the interpretation of each of the types of extended reported speech introduced in this section, but a common element in these examples appears to be that they all cast the reported speaker in a different role: as a thinker, as someone who holds an intention as described in the Report, or as a referent with specific attitudinal qualities.

Such cognitive activity appears to be completely absent in the next subtype, constituted by aspectual/temporal examples of

extended reported speech as in (2) above. Two further examples of this type are shown in (9).

- (9) a. KALMYK (Mongolic; Russia)
mōrən ūk-čə oɔə-n gi-ʒā-nā
 horse die-CVB.IPFV go.away-CVB.MOD say-PROG-PRS
 ‘The horse is close to death.’ (Baranova, 2015, 71)
 ME: The horse is saying: “I go away, I will die”
- b. WAN (Mande; Ivory Coast)
ké wà tí gé bā bō mō
 if time say LOG arrive PRT
 ‘When the time comes...’ (or ‘When the time is about to arrive.’) (Nikitina, 2020, 88)
 ME: If the time says: “I arrive”

Both examples in (9), like (2), have non-human subject referents in the Matrix unit, so it is clear that they do not involve actual speakers, but, more importantly, the Report unit describes an inceptive event that does seem to reflect any other perspective than that of the current speaker uttering these sentences. The converb constructions in (9a) occur in regular reported speech constructions and allow for a direct speech translation (Baranova, 2015, 64), but the example is not a statement about the mental state of the horse. Similarly, in (9b) the communicative relevance of the example is not some dramatic re-enactment of visions of time; it is about the inceptive or inchoative aspectual status of the content of the Report unit⁶.

Once more, it should be stressed that the inchoative interpretation in these examples is not a poetic invention by the speakers of these sentences. Rather, the examples represent a common way to express aspectual meanings in these languages. Birchall (2018) describes similar examples for languages of the Chapacuran family as expressions of incipient action or future tense. Güldemann (2008) also demonstrates the future tense meaning for other African languages and van der Voort (2002) reports it for Kwaza (see **Appendix Table A1**).

Another example in which the Report unit does clearly not signal an utterance or mental state is the NAME type, as in (10).

- (10) NGARINYIN (Worrorran; Australia)
malyan-nga junba jandu jirri ngarr-wəa2a2-ma-ə-nangka
 for.nothing-ONLY dance designer M.ANAPH IPL.INCL.S-IRR-do-PRS-3SG.IO
 ‘We don’t just call him ‘corroborree designer’ for nothing’ (Coate, 1966, 106, lines 39–40, cited by Spronck, 2015, 98, gloss updated)
 ME: We do not say “corroborree designer” about him for nothing

The term *junba jandu jirri* ‘the dance designer’ in (10) does not refer to a specific speech act, but is a general description of the oblique referent, which can be translated into English with the lexical verb ‘call’ or ‘name’. Among the examples of extended reported speech illustrated here, this type is slightly different in the sense that the ‘name Report’ is commonly assumed to be spoken, but the status of the

Report does not correspond to an utterance, which qualifies this as an extended meaning. Among the literature surveyed for this section, similar examples are attested in Ainu (Bugaeva, 2008) and in African (Güldemann, 2008), Tibeto-Burman (Saxena, 1988) and Siberian languages (Matić and Pakendorf, 2013).

A final type that we would like to introduce in this section is extended reported speech used for the purpose of information structuring, specifically topic marking, as in (11).

- (11) BEIJING MANDARIN (Sino-Tibetan; China)
wǒ jiāglái jiù shì shénme yā, shuō kǎo dàxué, ríguó kǎo-bu-shàng, wǒ jiù
 1SG future then be what PRT sayTOP sit:exam university if exam-NEG-DIR 1SG then
lián kǎo liǎng cì.
 successive sit:exam two CLF
 ‘So what will I do in the future?: as for taking university entrance exams, if I’m not admitted, I’ll just repeat the exams.’ (Chappell, 2012, 88, gloss updated)
 ME: {So what will I do in the future?} Say “take university entrance exams” {if I’m not admitted, I’ll just repeat the exams}

In examples like (11) the Report unit describes information that, presumably, has already been raised in the conversation and is subsequently commented on. Interestingly, information structuring examples of extended reported speech are described as signalling both that the content of the Report unit is a ‘topic’ and that the content is ‘highlighted’, which would rather suggest a focus function. Matić and Pakendorf (2013) also attest reported speech with a topic interpretation in their Siberian sample, and, more generally, discourse functions are attested in Aguaruna and African languages, as indicated in **Appendix Table A1**.

Güldemann (2008), 510 and Reesink (1993), 223 furthermore report that extended reported speech may have a ‘listing’ interpretation (e.g., ‘say *x*, say *y*, say *z*’), which could be seen as an instance in which the Reports are presented as a series of discourse topics.

2.2.2 Grammaticalised Extended Reported Speech

The interpretations of extended reported speech described in the previous section mostly corresponded to common reported speech constructions in the respective languages. They also shared the feature that the Matrix unit often corresponded to a lexical (matrix) verb in English, that is, ‘think’, ‘want’ or ‘call’, although the translations were more diverse for the attitudinal, aspectual/temporal and information structuring types of extended reported speech. All authors cited specifically introduce these examples because they represent common, conventional ways to express the meanings described and, hence, they involve a degree of constructionalisation. However, in most cases, the elements involved in the constructions do not appear to have developed into grammatical formatives.

This is different for the types that we will discuss in this section, for which a more straightforward argument can be made that the constructions have conventionalised to a degree that, at least in some languages, they have fully grammaticalised. A useful starting point for classifying these types is the overview in Kuteva et al. (2019), who list no fewer than eleven morphological categories into

⁶The observation that the Matrix unit in (9b) is subordinated under the conditional/temporal adverb *ké* ‘if’ is potentially relevant for the interpretation of this example, but not a requirement, as demonstrated by (2). We will explore potential connections between morphosyntactic structure and interpretation in **Section 4**.

which the lexeme SAY may grammaticalise⁷. These meanings/categories are shown in (12).

- (12) SAY (Kuteva et al., 2019, 375-388)
- CAUSE
 - COMPLEMENTIZER
 - CONDITIONAL
 - DISCOURSE MARKER
 - EVIDENTIAL, QUOTATIVE
 - EVIDENTIAL, REPORTED
 - PURPOSE
 - QUOTATIVE
 - SIMILATIVE
 - SUBORDINATOR
 - TOPIC

Since the classification developed by Kuteva et al. (2019) constitutes a 'lexicon of grammaticalisation', and grammaticalisation is defined as a diachronic process in which a lexeme becomes a grammatical element, presenting the types in (12) as deriving from SAY is a useful shorthand. Note, however, that as with the examples of extended reported speech presented before, the types of extended reported speech illustrated in this section commonly include a recognisable Matrix and Report unit. What characterises these types, though, is that, more frequently than in the previous examples, these units are integrated into other morphosyntactic structures. For this reason, 'grammaticalised' extended reported speech is often slightly distinct from other reported speech in the respective languages. The examples introduced here, therefore, often carry slightly more structural cues than those presented in the previous sections as to their 'extended' interpretation.

Taking the list in (12) as a guide, we will briefly illustrate the various types below. The CAUSE function (12a), exemplified in (13), appears to be particularly common.

- (13) a. KWAZA (Isolate; Brazil)
- 'ja kwi-'ni-da-ki*
already drink-CAUS-1SG-DECL
'I already let (him) drink (van der Voort, 2002, 320, gloss updated)
ME: I already said: 'let him drink!'
- b. ORO WIN (Chapacuran; Brazil)
- maw era n-on*
go.SG 2SG.FUT 3.NFUT-3M
'He sent him away' (Birchall, 2018, 21)
ME: "You will leave" he (said) to him
- c. NGARINYIN (Worrorran; Australia)
- wurrumiy-angga bud-ma-ra-ngarrugu*
w-class:ob.3pl.sub-fut.-'recognize' 3pl.sub-ma-past-1pl.incl.obl.
'They wanted us to recognize it' (Rumsey, 1990, 349, gloss updated)
ME: "They will recognize it" they said with regard to us

As the Mock English translations in (13) illustrate, each of these examples can still be interpreted as reported speech, so in this sense the function is less clearly grammaticalised than some of the other ones discussed below. However, the examples in (13) all involve an interpretation of (indirect) causation that partially requires a structural re-analysis of the reported speech

⁷This makes the lexeme SAY the most productive source for grammaticalised elements in Kuteva et al.'s lexicon, with only the entry 'locative' listing more functions.

construction involved: in (13b) and (13c) the entity who is coerced into performing the act described in the 'Report' is introduced as an oblique object in the Matrix unit. This involves a change in semantic roles: in both examples the subject of the Matrix becomes the 'causer' argument and in (13b) the indirect object, i.e. the 'addressee' is interpreted as a causee, as is the oblique object, i.e. the 'object talked about', in (13c). In (13a), the causal interpretation appears to arise slightly differently because of the presence of a morpheme glossed as causative in combination with the reported speech construction. In this example, the causee is left implicit.

These three examples already show that even though extended meanings in reported speech may be similar across languages, it should not necessarily be assumed that these meanings arise through exactly the same (diachronic) pathways.

Some typical examples of the complementiser function of reported speech (12b) are shown in (14).

- (14) a. SANZHI (Northeast Caucasian; Russia)
- T'ahir-ri-j han b-ič-ib [du q'an Ø-iy-ub Ø-ik'-ul]*
Tahir-OBL-DAT seem n-be.PFV-PRET 1sg late m-be.PFV-PRET m-say.IPFV-ICVB
'Tahir thought that I (=current speaker) was late.' (Forker, 2019, 185)
ME: It seemed to Tahir saying I was late
- b. AINU (Ainu; Japan)
- nep Ø=ihuy hum an sekor ihumar=an*
something 3.S=burn NONVIS.EV be.SG QUOT search=LOG.S
'We were searching (here: smelling around) whether something was burning.' (Bugaeva, 2008, 62)
ME: We search saying something appears to be burning

Judging by the Mock English translations of (14a) and (14b), these contain a redundant verb of saying that serves the main function of connecting a main clause describing some cognitive activity with a complement clause specifying this cognitive activity. Both the complementiser (12b) and the more general subordinator use (12j) of reported speech constructions are introduced more fully in Section 3.2.3, but the examples in (14) already reveal two important qualities of this subtype of extended reported speech. On the one hand, it is less obvious that these examples involve a Matrix and Report, since the 'complementiser' interpretation only emerges in the context of another bi-clausal structure. Therefore, two equally plausible analyses present themselves: either the Matrix and Report units fully overlap with these two clauses (e.g., the clause between square brackets in (14a) both derives from a Report and is a complement clause of the preceding clause *T'ahir-ri-j han b-ič-ib* 'it seemed to Tahir') or the verb SAY grammaticalises as a complementiser without bringing its associated Matrix and Report structure. We will briefly discuss this problem in Section 3.2.3, but refer to each of the examples cited here as extended reported speech. Second, even though we refer to the function in (14) as a complementiser, both examples represent cognitive actions, which raises the question of to what extent the 'SAY complementiser' interpretation can be generalised beyond predicates expressing meanings closely related to speech and thought. Matić and Pakendorf (2013), in particular, do show a variation of complement types with which a SAY-derived complementiser may occur: in some languages such an element may only combine with speech or cognition complements, in others it extends further, e.g. to verbs of perception and (eventually) any

complement/subordinate clause. The examples of reported speech ‘conjunctions’, listed in **Appendix Table A1**, may fall on various parts of this spectrum, and we will discuss these varying degrees of grammaticalisation in **Section 3.2.3** as well.

Another clause linking function, that seems related to the attitudinal senses illustrated in (8), is the conditional function (12c), as in example (15).

(15) GA (Niger–Congo; Ghana)

máha o níyeni ké oba.
give:1:SG:FUT you food (say) you:come

‘I’ll give you some food if/when you come.’ (Heine and Kuteva, 2002, 265, cited from Lord)
ME: I will give you food say you come

A generalisation that could be made over this subtype is that in (8) and (15) the ‘Report unit’ indicates a hypothetical or otherwise qualified event or action. As we will discuss in **Section 4**, this meaning can be derived quite simply from the full meaning of a reported speech construction, which, as we will argue, is also the case for the following three functions on the list: discourse markers (12d), as Chappell (2012) illustrates (and which might also include the ‘listing function’ referred to above) and the evidentials ‘quotative’ (12e) and reported (12f). The distinction between these evidential categories is variously defined in the literature: Aikhenvald (2004) suggests that quotative evidentials introduce a specific source referent (i.e. the reported speaker is explicitly mentioned), whereas reported/repor(ta)tive evidentials, otherwise labelled ‘hearsay’ or ‘reported evidence’, do not. However for Boye (2012) the relevant distinction lies in the semantic status of the Report unit: a reportative embeds a proposition, while a quotative embeds a speech act (also cf. Wiemer, 2018). Kuteva et al. (2019, 381) note the close diachronic relation between the two evidential categories.

The next function Kuteva et al. (2019) list is that of purpose (12g), cf. (16).

(16) BESLENEY KABARDIAN (Northwest Caucasian; North Caucasus)

[[qə-z-de-ʔapəqʷə-n] ʕʕʕʕ-ale-m zʕ-je-rjə] qə-kʷ-a
DIR-1SG.IO-COM-help-POT youth-OBL mouth-3SG.A+say-ADD came

‘The boy came to help me.’ (Ershova, 2012, 14)
ME: “Maybe help me (= the current speaker)” the boy saying, {he came}

The purpose interpretation appears on the one hand related to the WANT or intention interpretation as illustrated in **Section 2.2.1**, but the translation ‘in order to’ also reflects a more grammatical interpretation, which involves elements that may be used to introduce additional syntactic constituents. Like in the ‘complementiser’ examples, the Matrix unit in (16) occurs in subordination (the additive marker marks the Matrix as a converb Ershova, 2012, 76)⁸.

⁸The Matrix verb in (16) contains the incorporated noun ‘mouth’, which could suggest that the intention is actually spoken, but this construction is also used for the expression of thought (cf. Ershova, 2012, 78), so it does not seem a necessary interpretation for this example.

Example (16) is notable for another reason: the striking indexical features of the embedded first person pronoun, which refers to the current speaker, and not to the subject of the matrix clause. Such indexical patterns are in part a typical genetic property of languages like Beslenev Kabardian, but also hold implications for the relation between common categories of reported speech, such as direct and indirect speech in relation to extended reported speech. Unlike the impression sometimes given in the literature⁹, extended reported speech is not restricted to typical direct speech structures (as can also be seen from logophoric examples as in (2) and apparently indirect constructions, such as 15 and 16). For further discussion, see **Section 4.2**.

Purpose interpretations are common among the languages listed in **Appendix Table A1**, but an interesting further extension occurs in Tibeto-Burman (Saxena, 1988); the interpretation ‘to do intentionally, deliberately’, i.e., *on purpose*. Cf. (17).

(17) NEWARI (Sino-Tibetan; Nepal)

rām-ʃ gilās kurke dha-k-a-a kurk-l-ʃ
Ram-ERG glass break-NPC say-CAUS-PTCP-NF break-PD

‘Ram broke the glass deliberately.’ (Saxena, 1988, 381, gloss modified)
ME: {Ram broke the glass} saying making it “break”

The ‘on purpose’ meaning of (17) clearly constitutes a slightly separate type from the more common ‘purpose’ interpretation which Kuteva et al. (2019) distinguish, but like many of the other more grammaticalised examples of extended reported speech it too involves a subordinating structure, specifically a Matrix consisting of a participle predicate.

In addition to ‘evidential quotative’, Kuteva et al. (2019) also list a separate category of ‘quotative’, which refers to what Güldemann (2008) calls a ‘quotative index’: a Report unit that consists of a single morphological element that (often) diachronically derives from a lexical verb SAY. Although such Report units may develop extended meanings, they do not necessarily count as examples of extended reported speech under our definition in (4b).

We discuss the subordinator function (12j) together with complementation (12b) in **Section 3.2.3** and we have illustrated the information structuring subtype of ‘topic’ (12k) in 11 above, which leaves only one final class from Kuteva et al.’s list; that of similitive (12i), as illustrated in (18).

(18) TAMIL (Dravidian; India/Sri Lanka)

kumaar puli en-a paay-nt-aan.
Kumar tiger say-INF jump-PAST-3:M:SG

‘Kumar jumped like a tiger.’ (Heine and Kuteva, 2002, 269, cited from Lehmann)
ME: {Kumar to say tiger he jumped}

Comparison/similarity meanings are attested rather widely in the literature (cf. Güldemann, 2008; Matic and Pakendorf, 2013) and, like attitudinal meanings, can be derived from a common semantic component of reported speech constructions, as we argue in **Section 4.1**. Note that, as in many of the examples of

⁹For example, Pascual (2014), 83 presents her pioneering study as a ‘cross-linguistic study of *direct speech* for non-quotation’ (emphasis added), despite citing examples that do not represent direct speech in the chapter and allowing for a more inclusive description of the phenomenon elsewhere.

grammaticalised extended reported speech in this section (but not the causative subtype), the Matrix predicate *puli* ‘to say’ in (18) appears in a non-finite form.

2.3 Interpretations of Extended Reported Speech: An Inventory

Despite the wide variety of interpretations illustrated above, what stands out in the literature is how regular the meaning extensions in reported speech appear to be across unrelated languages. None of the subtypes illustrated in the previous sections appears only once in the literature summarised in **Appendix Table A1** and the very few additional functions that are attested can be related to more regularly described ones. For example, Saxena (1988) distinguishes ‘expletive’ and onomatopoeic functions in Tibeto-Burman, which indeed do not constitute typical Reported units, but may be categorised as a form of speech and/or sound emission.

One possible further subtype is mentioned by multiple sources but not included in Kuteva et al.’s (2019) list of grammaticalised functions. This is the category of ‘auxiliary’ and/or ‘light verb’, which Güldemann (2008) and Matic’ and Pakendorf (2013) find in their African and Siberian samples, respectively. This type reflects the observation that the verb SAY (or, more accurately, a predicate diachronically related to the meaning SAY) can bleach semantically over the course of grammaticalisation to the extent that it no longer has any distinguishable lexical meaning. As such, it often combines with types illustrated above, like the aspectual interpretations in (9) or the causative ones in (13). In such examples, the (historical) verb SAY does not contribute any lexical meaning to the construction, but merely connects elements in the sentence, or hosts cross-referential or temporal affixes, like a light verb (cf. Matic’ and Pakendorf, 2013, 385).

With respect to our present analysis, two aspects of this observation are relevant: on the one hand, first, it constitutes a rather different level of generalisation to the one adopted for most of the examples introduced above, that is, it focuses on the predicate SAY, rather than a full reported speech construction and, second, cross-linguistically, the development from speech verb into light verb can be seen to occur in the opposite direction in some languages. Particularly, for a number of Australian languages it has been observed that instead of having a specialised speech predicate, reported speech constructions in languages such as Ngarinyin (Rumsey, 1990) and languages of the Nyulnyulan family (McGregor, 2014) contain a generic action verb, often glossed as ‘do’ (cf. example 10). In the grammatical context of a reported speech construction this predicate assumes the lexical meaning ‘say’.

While assuming that the interpretations illustrated in the preceding sections arise out of grammaticalised (or re-lexicalised) uses of the lexeme SAY is a possible analysis for some languages, it is less appropriate for others. It is also variably applicable to the subtypes of extended reported speech so far introduced. For example, the complementising/linking function may be inviting focus on the word unit of SAY itself, but it equally involves a link between two clauses,

not unlike the Matrix and Report units already involved in a reported speech construction. If our analysis of meaning extension starts from a lexeme SAY, it is problematic to argue that the verbs used in (extended) reported speech may either entirely lose their speech interpretation, or that non-speech verbs can be recruited as matrix verbs in reported speech. This is not the case if we take reported speech constructions, i.e. Matrix and Report units with or without a lexical speech verb as the (diachronic) source for the extensions reported here.

This analysis also provides a consistent solution for the possible problem van der Voort (2002) diagnoses, that meaning extensions of the type illustrated in the preceding sections occur *regardless* of the lexico-grammatical status of the Matrix. Even affixes or particles like quotatives, or highly abstract constructions like the reported speech construction formed by the declarative marker in Kwaza (13a), may give rise to such interpretations as ‘want’ or ‘cause to do’. This creates the theoretical problem that under the SAY grammaticalisation analysis we would have a lexical meaning emerging from a grammatical construction (i.e., degrammaticalisation)¹⁰. Furthermore, simply focusing on the lexeme SAY removes from sight the similarities with meaning extensions arising from other types of Matrix units.

Before exploring the consequences of this integrated approach to extended reported speech further, let us take stock. The observations in **Section 2** expand the initial inventory of extended functions of reported speech based on Pascual (2014) in (3) to the set of functions in (19). Although the distinction between lexical and ‘grammatical’ functions is not clear-cut, we may further divide these functions into a more lexical group summarised in (19a) and a group that bears a resemblance with morphosyntactic categories, or functional elements in the sentence, listed in (19b).

- (19) a. *lexical extensions*
 NAMING: call;
 THOUGHT: think, want.
 b. *morphosyntactic extensions*
 ATTITUDE: modality, irrealis;
 EVIDENTIALITY: quotative, reportative;
 TIME: inchoative aspect, future tense;
 VALENCY CHANGING: causation, reason, purpose;
 CLAUSE LINKING: complementiser, connective;
 INFORMATION STRUCTURE: topic, highlighting.

Before placing the functions in (19) in a broader context in **Section 4**, we will first try to delve slightly deeper into the distribution and origin of some of these functions, by presenting a typological study of two specific subtypes of extended reported speech in **Section 3**. As we will show, there are many difficulties inherent in studying extended reported speech as a typological topic, but in order to contextualise the

¹⁰Depending on an author’s theoretical stance, this situation may or may not jeopardise their account, but in any case it complicates it if one has to unify the observation that a similar meaning extension arises from two different sources (a lexical and a non-lexical one), which is an additional step not required for the analysis that the Matrix unit is the relevant source element.

observations above it will be useful to gain an impression of how widespread the phenomenon is in the languages of the world. In order to develop an understanding of how extended meanings arise out of the structural features of reported speech constructions, we will also present brief case studies of two such meanings, that is, the WANT and complementiser/linker subtypes, which can be identified relatively reliably in descriptive grammars.

3 A SAMPLE STUDY

3.1 Methodology and Distributions

In this section we present the first results of a broad typological study on extended reported speech based on a cross-linguistic, genetically balanced sample of 100 languages. We study the distribution of the phenomenon, aiming to show that it is not restricted to certain areas or language groups but can be found around the world (**Section 3.2.1**) and present case studies of extended reported speech with a WANT interpretation (discussed in **Section 3.2.2**) and with a complementising/clause linking function (see **Section 3.2.3**). The purpose of these case studies is to examine structural similarities between examples of extended reported speech with comparable interpretations in unrelated languages, which should lend insight into how these interpretations arise. The two subtypes chosen are particularly useful for such an exploratory analysis, since we will be able to draw on some clear hypotheses for such structural features based on previous literature, which we will be able to test on the basis of our sample.

Before presenting these results, however, we introduce our sample and sampling procedure in **Section 3.1.1** and briefly reflect on our methodology and its possibilities and limitations in **Section 3.1.2**.

3.1.1 Sample

Linguistic typology is a branch of linguistics that seeks to classify and understand the range of variation found in the world's estimated 7,000 languages. It does so by conducting sample studies of features that are explicitly pre-defined on the basis of semantic and/or abstract formal properties (Haspelmath, 2010), mostly using descriptive grammars, i.e., maximally comprehensive descriptions of individual languages organised in a way that allows for cross-linguistic comparison.

The selection of languages in a typological sample, Rijkhoff et al. (1993) suggest, qualifies these samples as one of two kinds: probability and variety samples. Probability samples are intended as a maximally representative selection of the world's languages, aimed at answering statistical questions about the frequency with which a feature occurs. To this end, larger language families are better represented in probability samples than smaller language families and the primary focus is on diffused categories¹¹. Variety samples, on the other hand, aim to capture a maximum amount of genealogically and topographically distinct languages. To this end, larger language

families are not prioritised over smaller ones in the sample, which means that typologically 'rare' languages are included in the same ratio as more familiar ones. A variety sample allows us to address the qualitative question whether a linguistic feature is restricted to a particular area or language group and within what range the observed values fall.

For our purposes of demonstrating that extended reported speech (as defined in 1) occurs globally and to understand the variability of the phenomenon, our case study involves a variety sample, constructed following the method proposed by Miestamo et al. (2016). This method is based on the distribution of languages across six macro-areas and according to a classification in genera, defined by Dryer (1989) as a set of closely related languages with a common time-depth of no more than 3,500 to 4,000 years. Such a classification is inherently subject to ongoing academic debate, with occasional reclassification of individual genera as new diachronic evidence emerges, but for our sample we follow the list of genera distinguished in Dryer and Haspelmath (2013). The notion of genus also allows us to take into account the diachronic influence of language contact in areas where genetically diverse languages have long been in close proximity, which could indicate patterns of borrowing.

In constructing our sample, we have randomly selected 100 genera, following the areal distributions proposed by Miestamo et al. (2016), but have favoured languages with larger descriptive grammars over languages with fewer available resources in order to maximise the chance of finding relevant descriptions of extended reported speech. The full sample of languages, including the respective genera and sources used is described in **Appendix Table B1**.

3.1.2 Methodological Limitations: What This Study Can and Cannot Tell Us

A typological study as attempted in this section faces the obvious challenge that negative evidence does not demonstrate non-existence and positive evidence is not necessarily exhaustive. Put differently, if a descriptive grammar does not present examples of extended reported speech in accordance with our definition this cannot be taken as evidence that the phenomenon is absent in the respective language and if a descriptive grammar does include examples of extended reported speech, these do necessarily illustrate the full range of functions that the phenomenon can have. Unlike the specialised studies surveyed in **Section 2**, the descriptive grammars examined here do not aim to provide a full and detailed account of extended reported speech and may be based on corpora that lack the phenomenon, even though it exists in the language concerned. For each of the languages in our sample, we fully rely on the judgements by the author of the grammar, who, no matter how thorough and comprehensive the description, inevitably presents a 'doculect' (Cysouw and Good, 2013), a language-as-described based on a limited amount of contexts of use and selected, glossed and analysed by an author. Therefore, distributions may under-represent occurrences of extended reported speech if the corpora on which a description is based did not include them, even though extended reported speech does occur in the language. On the other hand, accounts of extended reported speech may be relatively over-represented in languages that belong to an area in which extended reported speech posited is as an areal feature (e.g., Cohen et al., 2002) so that it is on the radar of the respective grammar writer.

¹¹An increased awareness of the importance of language contact and Sprachbund phenomena in the spread of linguistic features casts doubt on the assumption that genealogical affiliation can be taken as a primary selection criterion in probability samples, but this issue should not concern us here.

Despite these limitations, using the definition of extended reported speech in (1) we should be able to identify relevant examples in the sample. We should not expect the phenomenon to be limited to any specific area and to only involve a specific number of meanings. We would also not expect the phenomenon to be limited to certain structural types of reported speech, or involve any particular grammatical features. However any patterns we do find will lend further insight into the nature of extended reported speech.

In this section we only explore a few such patterns with respect to two subtypes of extended reported speech, but for a fuller analysis of the sample see Casartelli (fc). We begin with a more general question: where can examples of the phenomenon be found?

3.2 Results

3.2.1 Distribution

The map in **Figure 1**, based on the specific studies surveyed in **Section 2**, suggested that extended reported speech is not an isolated phenomenon only attested in some parts of the world, but occurs independent of language families or contact areas. The 100-language sample affirms this impression, indicating that we find relevant examples on all major continents.

Figure 3 shows the distribution of such examples: for the locations of languages indicated in orange we find evidence for the occurrence of extended reported speech in accordance with our definition in (1), for the ones indicated in blue the respective descriptive grammars do not include such examples¹². As discussed, these observations cannot be taken as definite proof that extended reported speech is absent from the respective language, just that in the most comprehensive description of this language it has not been raised as an example or theme.

Figure 3 does not specify the types of meaning extensions found in the sample. For a fuller analysis the reader is referred to Casartelli (fc). However, the distribution confirms the wide spread of extended reported speech across areas and language families, with about half of the languages in the sample displaying the phenomenon (see **Appendix Table B1** for a list of included languages).

The discontinuities in the distributions in **Figure 3** are somewhat more instructive than the continuous groups of blue or orange dots, since our main goal is to demonstrate the occurrence of extended reported speech independent from geographical regions. Nevertheless, two areas slightly stand out: the sample does not include instances of extended reported speech in the languages of Western Europe, whereas in South-East Asia sources quite commonly describe it. Although such patterns should be interpreted with care given the considerations discussed above, they highlight the distinction between our more restricted notion of extended reported speech, as opposed to the common phenomenon of the creative, metaphorical use of

conversation to express non-speech meanings in fictive interaction (Pascual, 2014). While the latter forms of use are common in (spoken) Standard Average European languages, extended reported speech is not¹³. This is particularly clear in the case of Catalan, which figures prominently in the literature on fictive interaction with examples such as (3).

(20) CATALAN (Indo-European; Spain/Andorra/France/Italy)

Ara digue'm tu a qui se li acudiria fer una cosa així.
now tell-me you to who it occur-would to do something like that

'You tell me who would do something like that' (Pascual, 2002, 181, gloss and translation as in the source)

Example (3) counts as fictive interaction since the addressee of this utterance is not actually expected to tell anything about the person 'who would do something like that', but it is not an example of extended reported speech within the definition provided in (1). This is not to say that such examples definitely do not exist in Catalan or any of the other SAE languages in our sample¹⁴: as indicated in **Section 3.1.2**, it simply means that using the selection criteria we have set for our study we have not identified such examples in the descriptive grammars.

Although the more general cognitive principles that Pascual (2014) describes are likely to be relevant for both synchronically metaphorical uses of fictive interaction and lexicalised and grammaticalised forms of extended reported speech, our approach visualises the latter phenomenon and shows that it can be demonstrated to occur relatively frequently around the world.

3.2.2 WANT

In this section and **Section 3.2.3** we will illustrate two different subtypes of extended reported speech in our sample: examples with an intention/WANT interpretation and those with a complementiser interpretation. Our aim with these case studies is to examine an aspect of the phenomenon that has so far received little attention, but that has important implications for our understanding of extended reported speech in relation to perspective expressions more widely and other types of reported speech in particular. This concerns the (diachronic) structural means through which the relevant meaning extensions arise.

Our reason for focusing on these two subtypes, the 'lexicalised' interpretation WANT and the 'grammaticalised' complementiser subtype, is that for these two classes of examples the literature presents sufficient evidence to form hypotheses about cross-

¹²Like the map in **Figures 1, 3**, was produced using the R-package *lingtypology* (Moroz, 2017).

¹³It is likely that European sign languages show more evidence of extended reported speech, given other observations about grammaticalised forms of fictive interaction found by, e.g., Jarque and Pascual (2015) and Jarque (2016). Unfortunately, our sample only includes oral languages but the increasing availability of descriptive grammars will hopefully allow us to discuss examples from sign languages in future work.

¹⁴And this English sentence is, in fact, an indication that fictive interaction is a much broader phenomenon than extended reported speech.

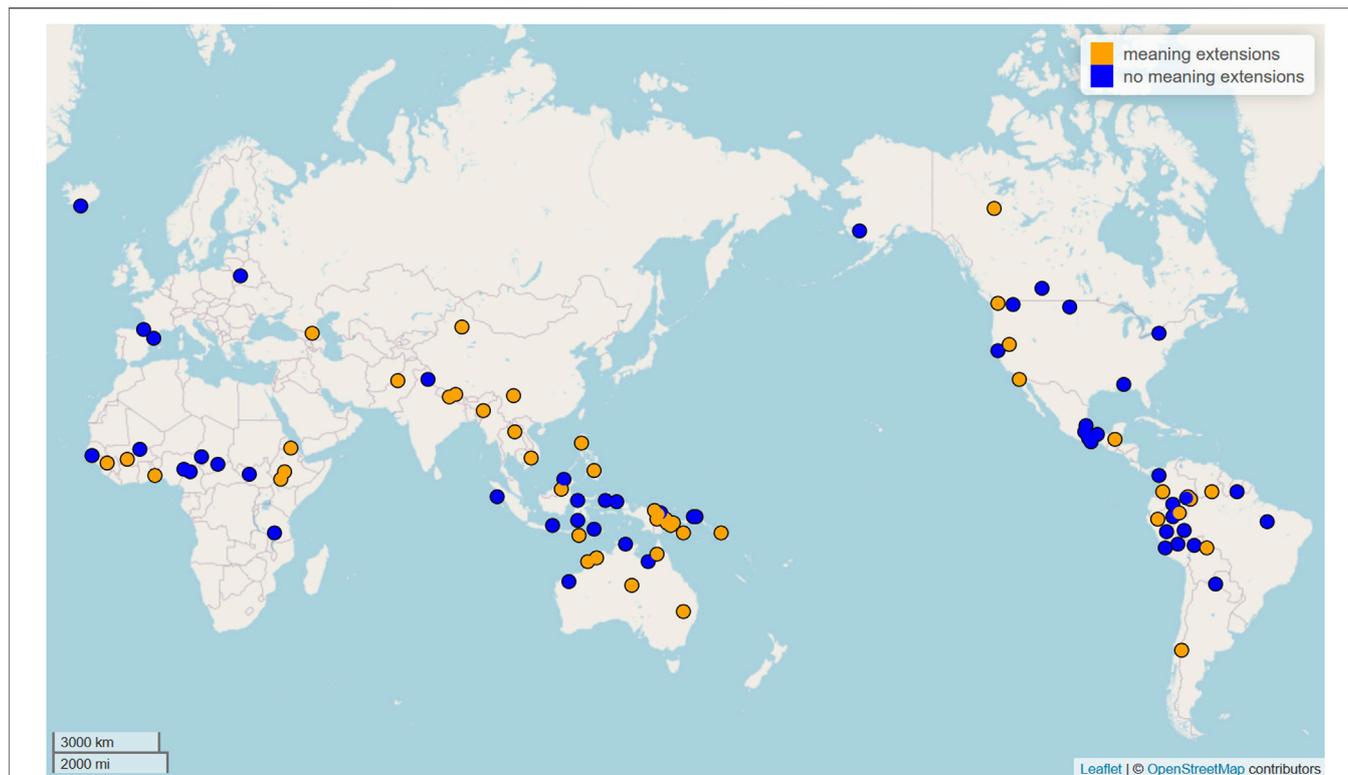


FIGURE 3 | Extended meanings of reported speech in a 100-language sample.

linguistic regularities in their structural composition¹⁵. The WANT interpretation of extended reported speech has variously been described as an ‘intentional’ (cf. Everett, 2008; Konnerth, 2020) or ‘desiderative’ (cf. McGregor, 2007) construction but its cross-linguistic structural realisation appears to be rather consistent: as first described by Rumsey (1982) for Ngarinyin, it often includes an embedded first person and a non-present/non-actual tense in the Report. The schematic representation in (21), adapted from Spronck (2015, 100), illustrates these features.

(21) [[... S_{1sg} + FUT]_{Report} SAY]_{Matrix}

Throughout this section we introduce various schematic representations of extended reported speech as in (21). Here and below, the order of the Matrix and Report elements is non-iconic: the representation in (21) may reflect a structure in which the Matrix either follows or precedes the Report. The order of the morphemes and lexeme SAY is variable as well. What is relevant, in this instance, are the person and number features of the subject and the future tense in the Report. Examples closely resembling the representation in (21) indeed occur relatively frequently in the sample in

extended reported speech with a WANT interpretation, as illustrated in (22).

- (22) a. DARAI (Indo-European, Eurasia)¹⁶
her nə məi punu səti dze-m kəhə-lə tə poran nihidzə dza-inei
 look-IMP PART I also immolation go-HOR say-PST PART life NEG go-MIR
 ‘Look, I want to immolate but I did not die!’ (Dhakal, 2012, 79)
 ME: {Look} “Let I also go immolation” I said {but life did not go}
- b. KALAM (Trans-New-Guinea, Papunesia)
“Laplap d-in!” ag ny-i ktg ow-a-k
 clothes take-1SG.HORT say think-SS.PRIOR leave come-3SG-PAST
 ‘She wanted to take (steal) clothes (but) left (them) behind.’ (Pawley and Bulmer, 2011, 64)
 ME: “Let me take clothes!” say, having thought, {leaving she came}.
- c. GUMER (Afro-Asiatic, Africa)
asa ə-t ‘əft̪ bar-əc-im mərəβ t̪ ‘əβat ‘əc-im
 fish 1s-take.IPFV say.PFV-3sIS-CV,M net take.PFV-3sIS-M
 ‘Wanting to catch fish she took a net.’ (Völlmin, 2017, 171)
 ME: {She took a net} saying: “I catch fish”
- d. WORRORRA (Worrorran; Australia)
nga-nya=ya kuN-nga=yi-eerri
 1-OPT=go VCOMP-1=do-PROG
 ‘I want to go.’ (Clendon, 2014, 351)
 ME: “Let me go” that I am saying/doing¹⁷

In addition to singular first person subjects in the Report, all examples in (22) are combined with a non-present tense or non-realis mood. Future tense occurs in several examples

¹⁵See Casartelli (fc) for more detailed analyses and accounts of other subtypes of extended reported speech.

¹⁶For the remaining examples in this section we list the macro-area in the sample, rather than countries in which the respective language is spoken.

¹⁷Like in other Worrorran languages (Rumsey, 1990) and Nyulnyulan languages (McGregor, 2014), the Matrix predicate *yi-in* in Worrorra can both be translated as ‘say’, ‘think’ or ‘do’. Clendon (2014) opts for the gloss ‘do’, but the description in the grammar demonstrates that ‘say’ is one of the available translations, qualifying this example as extended reported speech in accordance with our definition in (4b).

below (cf. 24b and 24c), but in these examples we find hortative or optative mood (22a, 22b, 22d), or imperfective aspect (22c). On the basis of these observations we may conclude that the future tense in the Report is slightly too specific: although it occurs in the sample, the common feature between all tenses and moods in the Reports of extended reported speech illustrated so far appears to be that they place the event described in the Report in some time other than the here-and-now. We will label this observation IRRealis, as in (23).

(23) [[... S_{1sg} + IRR]_{Report} SAY]_{Matrix}

In addition to first person singular, we also find other person and number values in the Reports of WANT extended reported speech, such as non-singular forms. In the Yeri example in (24a), both the subject of the Matrix and Report are first person plural. In (24b) the Matrix subject is coreferential with a first person dual in the Report. In contrast, (24a) has a third person subject in the Report and also in the Matrix. In accordance with (23), the tense/mood values in the Report units in (24) are all non-present/non-realis.

(24) a. YERI (Torricelli; Papunesia)

Ø-nobia h-ero.
1pl-talk.REALIS 1pl-go.IRR
'We wanted to go.' (Wilson, 2017, 380, gloss updated)
ME: We talked "we might go"

b. MA MANDA (Trans-New-Guinea; Papunesia)

sisa gaamiyong=li laai ku-ntaa-mot taa-ka ku-gu-mot
±2days PN=COM PN go-FUT-1DU say-SS go-RP-1DU
'The day before yesterday I wanted to go to Lae with Gamiyong, so we went.' (Pennington, 2018, 183)
ME: The day before yesterday I said "We will go to Lae with Gamiyong", {so we went}

c. OKSAPMIN (Oksapmin; Papunesia)

niu jia max boten x-t-pe=li-m xe-ja
new(English) year(English) RECG pray(Tok Pisin) DO-PFV-IF.PL=QUOT say-SEQ
be-PRS.PL
ME: They wanted to pray for, you know, that New Year.' (Loughnane, 2009, 124)
ME: They were saying like they will pray for New Year{, you know}

These examples indicate that rather than taking the specific person and number values first person singular as a typical feature of WANT extended reported speech, a better generalisation is to highlight what it signals: a first person subject in the Report necessarily indicates co-referentiality with the subject of the Matrix. In addition to first person singular marking in the Report, co-referentiality may also be indicated by having the same person/number values in both the Report and Matrix units, viz. first person plural in (24a) and in (24b) (also combined with same subject marking in the Matrix) and co-referential third person plural marking in (24c).

In accordance with these observations, we may update the schematic representation of WANT extended reported speech as in (25), in which the coreferential relations between the subject S in the Report and in the Matrix are indicated by the subscript index *i*.

(25) [[... S_i + IRR]_{Report} S_i SAY]_{Matrix}

The remaining examples of WANT interpretations in the sample show minor variations on the pattern illustrated above. Kambera in example (26a) has a first person subject in the report, but no apparent tense/mood marking on the

auxiliary verb 'try' (but note the ME based on the author's alternative translation with 'let's'). A similar observation can be made for the Paiute example in (26b), which has a generic tense (TNS) form. This form is due, however, to a morphosyntactic restriction in the language, which disallows the combination of any other tense forms with applicative marking (Thornes, 2003, 398).

(26) a. KAMBERA (Austronesian, Papunesia)

"Ka ta-kama wà-nyaj, [i Ama], wà-mu ní
CNI 1pl.NOM-try use-3sg.DAT ART father say-2sg.GEN DEI
'You want(ed) to try out father' (Klamer, 1998, 294, gloss updated)
ME: "Let's try out father" you said"¹⁸

b. PAIUTE (Uto-Aztecan; North America)

oo=kaina mi=tiya-pl-miu ta na-ni-naka-ki-ti
so=MOD PL=die-PFV QUOT 1.DU MM-IP/speech-hear-APPL-TNS
'Perhaps those who have passed on want us to hear them.' (Thornes, 2003, 340, gloss updated)
ME: Perhaps those who have died say about us (like) they themselves speak-hear with respect to us

Even though both examples in (26) could be seen as slight variations of the representation in (25), it appears to capture most of the examples of the WANT subtype of extended reported speech in the sample, and the previous literature (again, note that the word order in (25) is non-iconic).

This leads us to three preliminary conclusions: first, the relative similarity of WANT extended reported speech across unrelated languages and areas is unlikely to be coincidental. This suggests a more fundamental common factor underlying these examples. Second, the similarities between the occurrences of extended reported speech are not only semantic, the examples in this section also appear to share a structural basis. This observation is not new, for example Reesink (1993), 223 notes that all examples of extended reported speech in Usan involve a same subject marker, indicating co-referentiality between the subject of the Report and that of the Matrix clause, whereas 'regular' reported speech in the language does not require this. While we would not predict that all extended reported speech across languages can be qualified in terms of a restricted set of formal features, the relative frequency and correspondence of structures involved in extended reported speech deserves more attention than it has received in the literature so far. Third, and perhaps most significantly, the features as represented in (25) cross-cut common subtypes of reported speech, such as the binary opposition between direct and indirect speech. This has implications for our understanding of the boundaries between perspective constructions and non-perspectival constructions, as we argue in Section 4.

3.2.3 Complementiser/Clause Linking

Examples of extended reported speech displaying a complementiser/clause linking function are slightly less numerous in our cross-linguistic sample, but nonetheless occur five times across five language families and two linguistic macroareas¹⁹.

¹⁸This alternative translation is provided in the original source.

¹⁹Again, note that no conclusions can be drawn about the absolute or relative occurrence of this subtype of extended reported speech on the basis of these frequencies, since the sources in the sample do not necessarily provide a fully comprehensive overview of the phenomenon in the respective language.

Typical examples of this strategy are shown in (27), where in Gumer the construction ‘consists of a quoted sentence concluded by a converbial form of *bar* [‘say’] followed by the matrix verb’ (Völlmin, 2017, 168) and in Stiang, spoken in Cambodia and Vietnam, the clause linking function is expressed with a conjunctive form of the speech verb.

- (27) a. GUMER (Afro-Asiatic; Africa)

cənə-x^w-im ba-x^w-im od-x^w-in-im.
 come.PFV-1sgS·M' say.PFV-1sgS-CVB·M' tell.PFV-1sgS-3sgmO·M'
 ‘I told him that I came.’ (Völlmin, 2017, 168, gloss updated)
 ME: I told him saying “I came”
- b. STIANG (Austro-Asiatic; Eurasia)

jun praj paj gət [lah jəmra:h] paj ?alu: ?atəp hej
 deer running 3 know [say:CONJV cliff] 3 stop fall 1SG
 ‘The deer runs, he knows there’s a cliff, he stops and makes me fall.’²⁰ (Bon, 2014, 488, gloss updated and translated)
 ME: The deer knows saying “a cliff” {it stops and makes me fall}

In both of the languages in (27), the complementising use of speech verbs is restricted to the semantic domains of speech and cognition (Völlmin, 2017, 168; Bon, 2014, 487): in (27a) it precedes a form of *od-* ‘tell’, in (27b) of *gət* ‘know’.

A variation on the subordinated forms in (27) is shown in Darai (28), where the complementising speech verb receives a sequential marker.

- (28) DARAI (Indo-European; Eurasia)

lə dze^hi dul-hi-r-ke rə kanc^hi dul-hi-r-ke ijo kəh-ikun kəhə-lə
 PART elder wife-2SGPOSS-DAT and young wife-2POSS-DAT this say-SEQ say-PST
 ‘Well, (the hermit) said that that was for his elder wife and another (that) for his younger wife.’
 (Dhakal, 2012, 144)
 ME: He said saying this for your young wife and for your elder wife

Alternatively, SAY-derived ‘complementisers’ may also remain uninflected, as in (29).

- (29) a. FONGBE (Niger-Congo; Africa)

Siká / Kókú q̄ q̄ yé t̄m q̄ émi fin Ásibá s̄n gb̄s
 Sika / Koku say say 3pl know say LOG steal Asiba GEN goat
 a. ‘Sika; said that they know that she; stole Asiba’s goat.’
 b. ‘Koku; said that they know that he; stole Asiba’s goat.’
 c. ‘Sika/Koku said that they; know that they; stole Asiba’s goat.’ (Lefebvre and Brousseau, 2002, 79)
 ME: Koku say say they know say s/he/they stole Asiba’s goat
- b. KAKABE (Mande; Africa)

fóó káà kən kó kàrà pɛsɛɔ máá dí
 UNIV PFV.TR 3SG know say studies-ART get PFV.NEG be.easy
 Everybody knows that studying is not easy. (Vydrina, 2017, 151)
 ME: Everybody knows say the studies are not easy

The Fongbe example (29a) has three occurrences of *q̄* ‘say’, the latter two of which act as a linking element between the main and complement clauses, and to which we therefore refer as a complementiser. A similar structure is attested in (29b).

The examples above share several features that we might represent schematically as in (30): all include two clausal units linked by a non-finite form of SAY. As in the representation in

(25), the order of the elements in this representation varies depending on the respective language.

- (30) [[clause] SAY:non-finite [clause]]

The variable ‘SAY:non-finite’ may either constitute a non-inflecting form or a subordinate form of SAY in the examples given here, which seem mostly representative for the type of examples commonly presented for the complementiser type of extended reported speech adduced in the literature (cf. Klammer, 2000; Heine and Kuteva, 2002; Güldemann, 2008; Matic and Pakendorf, 2013; Kuteva et al., 2019). The examples all contain a clause on each side of the SAY verb, which is consistent with Matrix and Report units. However, they also contain an additional main verb, providing lexical meaning to the sentence/respective clause. For this reason, the units represented in (30) have received the more abstract label ‘clause’, although they could mostly be interpreted as (originating from) Matrix and Report units as well.

We would like to address three observations about the examples of the ‘complementiser’ subtype illustrated in this section and in Section 2.2.2. A first observation that stands out, particularly given the broad grammatical label ‘complementiser’ that we have given to this subclass, is the very small lexical range of main verbs with which it appears to combine: the lexical matrix verbs used in the examples above are either speech verbs as well (27a, 28 and 29a), or cognition verbs: more specifically, verbs of knowing (27a and 29b). In Section 2.2.2, examples (14a) and (14b) also involved cognition verbs, viz. ‘think’ and ‘search’, respectively. Consequently, calling the SAY:non-finite form in (30) a ‘complementiser’, ‘linker’ or ‘subordinator’ is perhaps slightly deceptive: in many languages, the application of this form is limited to only a small class of complement-taking verbs, closely related to the semantic domain of speech and thought.

This impression is affirmed by the complement types Heine and Kuteva (2002), 261–265/Kuteva et al. (2019), 375–379 and Matic and Pakendorf (2013), 372–375 illustrate, which mainly involve main clauses of speech, thought and knowing, as well as perception and fear. However, the gradual dissemination of the structure represented in (30) with various types of main verbs is instructive. On the one hand, it illustrates a common process in grammaticalisation, in which the shift from a lexical to a grammatical element is not a matter of all-or-nothing, but spreads from one or a few lexical combinations and constructions to ever more lexical contexts (De Smet, 2012). It also neatly suggests a path through which structures as in (30) become established, from occurring with more speech-like Matrix/main clauses, to increasingly less speech-related ones.

This suggest that qualifying the status of SAY:non-finite in (30) in strict categorial terms, i.e. as either a lexical element or a complementiser/‘linking element’, may not always be possible, because this status varies between occurrences. The types of non-finite forms found in the examples above reflect this as well: dependent inflections as in (27) may signal varying degrees of conventionalisation.

Studying the behaviour of complementiser uses of SAY in Austronesian languages, Klammer (2000) presents a similar conclusion about the syntactic status of these elements and

²⁰Original translation: ‘Le cerf court, il sait qu’il y a une falaise, il s’arr’te et me fait tomber’.

proposes that the interpretation falls out from the defective inflection patterns of SAY:non-finite forms which (30) displays. Specifically, the 'SAY complementiser' in languages that have it, commonly shows no or non-matching person features to co-index arguments in the main clause, and this 'bleached' argument structure coerces a 'bleached' semantic interpretation. Klamer's analysis is consistent with our findings and leads us to a second observation about complementiser extended reported speech: although the details differ, the WANT and complementiser subtypes of extended reported speech both have consistent structural features, as we have schematically represented in (25) and (30) that correlate with their respective interpretation. In both instances, these involve, among other features, the use or lack of certain person referential features and/or tense and mood forms²¹.

A third observation we would like to address here goes back to the complement types found in (30), which, in the sample, divide into speech complements and knowledge complements. The exact syntactic status of these complement types requires closer investigation for each individual language and would be weighed differently by various syntactic models, so we will refrain from detailed generalisations about the syntactic structures involved. However, there is widespread agreement among both formal and functional approaches to syntax that the scope properties of speech complements and knowledge complements (i.e., a clausal structure that expresses what-is-said as opposed to a clausal structure that expresses what-is-known) are distinct (Hooper and Thompson, 1973; Boye, 2012; Gentens, 2020). Specifically, the syntactic integration of knowledge complements is assumed to be 'tighter' than that of speech complements, which has direct consequences for their interpretation: the content of the former is asserted by the speaker, whereas that of the latter is not, cf. (31)²².

- (31) a. Alex said that [you knew that the batteries were flat.]*illocution*
 b. Alex knew that [you said that the batteries were flat.]*proposition*

As (31) illustrates, both orders of complement taking predicates are possible, but in (31a) the interpretation of the unit between square brackets is an illocution, some utterance attributed to Alex, which the current speaker does not state as fact. In (31b), the complement clause marked by the square brackets is asserted by the speaker: the suggestion that the second person referent actually made the statement about the batteries is an integral part of the speaker's message. This effect cannot simply be attributed to the difference between the verbal predicates 'say' and 'know' (see Gentens, 2020). It relates to more general observations about scope relations in language in which elements to illocutionary meaning have a wider scope and are less tightly integrated in clauses than, e.g., elements that relate

to epistemic meanings, which, in turn, have a wider scope than, e.g., temporal elements, cf. (32).

- (32) a. Alex, frankly, probably did not read the instructions quickly.
 b. [?]Alex quickly, probably, did not read the instructions frankly.

As has been observed by both functionalist and formalist grammarians (cf. Dik, 1997a; Dik, 1997b; Cinque, 1999), adverbs targeting various parts of a sentence can be used to diagnose boundaries and scope relations between them. In (32a), the adverb 'quickly' (a temporal adverb) only has scope over the activity 'read the instructions', 'probably' (an epistemic adverb) over 'did not read the instructions quickly' and 'frankly' (an illocutionary adverb) the entire sentence. Re-ordering the adverbs in (32a) with the effect that, e.g., temporal and epistemic adverbs have scope over an illocutionary adverb results in an unintelligible sentence (32b).

Readers will weigh observations like those about the English sentences in (31) and (32) differently and, depending on other assumptions about the nature of language, explanations vary. However, the idea that sentence units have distinct scope properties that correlate with their meaning and can be classified into units that are more and less deeply syntactically embedded, is both pervasive and robust (Hengeveld, 1989; Boye, 2012; Cinque, 2013).

With respect to the distinction between the complements found with the complementiser subtype of extended reported speech, we suggest that they seem to either constitute illocutions or propositions, which suggests varying degrees of syntactic integration (as in 31).

3.3 Summary

In **Section 3** we reported on the first results of a sample study into extended reported speech. All observations introduced here will be discussed further in Casartelli (fc), but the initial analysis revealed several properties of extended reported speech that provide further insight into the phenomenon.

We found that both of the subtypes examined that they display considerable structural similarities within each respective type. We also identified three more general processes in the grammaticalisation and conventionalisation of these subtypes, which we would like to summarise as in (33).

- (33) a. Recasting: the subject of the Matrix unit takes on a different role, e.g., from reported speaker to a reported thinker or a person with some intention;
 b. Rescaling: units within extended reported speech become associated with different scope properties, e.g., the difference between speech and knowledge complements;
 c. Bleaching: a more general process in grammaticalisation in which elements gain a broader semantic interpretation, which may be signalled by a reduction of structural complexity (cf. Klamer, 2000)

In the next section we will relate the three processes described in (33) to properties of extended reported speech more widely.

4 DISCUSSION: REPORTED SPEECH AND THE EVOLUTION OF GRAMMAR

In this section we place the empirical observations from the preceding sections in a broader perspective and suggest some implications for our understanding of reported speech as a linguistic structure and its relation to grammatical and lexical

²¹For similar observations regarding the Biblical Hebrew complementiser *lemor* and further analysis in the context of fictive interaction, see Sandler and Pascual (2019).

²²We thank a reviewer for emphasising the relevance of assertion for the interpretation of extended reported speech and apply it in our notion of 'rescaling' below.

meaning. First, in **Section 4.1**, we return to the three processes summarised in (33) and examine their role in the grammaticalisation of extended reported speech. Particularly, we relate these processes to the meaning of reported speech as a source construction for all the various meanings and structures observed in the preceding sections. In **Section 4.2** we briefly contemplate the variety of structures involved in extended reported speech and compare these to standard, commonly recognised subtypes of reported speech, particularly direct and indirect speech constructions and quotative/reportative evidentiality. We suggest that the observation that extended reported speech cross-cuts such classical categorisations of reported speech indicates that there is more continuity within the domain of reported speech than is sometimes assumed. Finally, in **Section 4.3** we return to the research programme of fictive interaction and propose an interpretation of extended reported speech that not only places metaphors of communication centrally in the way in which humans think and speak about the world, but that acknowledges meta-linguistic reflection and reported speech as shaping forces in the emergence and evolution of grammatical categories. This is, admittedly, a speculative story, but for us it is also a significant motivation for the importance of understanding the nature and variation of extended reported speech.

4.1 Back to the Source (Construction): Recasting, Rescaling and Semantic Bleaching in Extended Reported Speech

The analysis that (at least some of) the meanings attested in extended reported speech fall out from a diachronic process of semantic bleaching suggests that it should be possible to relate them to meaning components in the original source construction. Spronck and Nikitina (2019) propose three such meaning components for reported speech, as summarised in (4.1):

- (34) a. *'A semiotic status of 'demonstratedness'*
This component qualifies the reported message as a representation of an utterance, as demonstrated (or 'depicted') 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' (Spronck and Nikitina, 2019, 25).

Despite the great variety of forms of reported speech in the languages of the world, the definition in (34) suggests that a reported speech construction should at least indicate three meanings: first, it should signal, as per (34a), that the Report unit is 'demonstrated' or 'depicted' rather than stated (Clark and Gerrig, 1990; Davidson, 2015; Clark, 2016). This property sets R apart from immediately surrounding clauses. As per (34b), reported speech also introduces an opposition between a perception event and the current speech event, which is the definition of an evidential meaning Jakobson (1957) coins. Third, as per (34c), reported speech explicitly or implicitly allows for (inferences about) the attitude of the current

speaker towards the content of the attributed utterance (cf. 'distancing' in terms of Guldemann, 2008)²³.

If the definition in (34) is on the right track, the process of 'semantic bleaching' in extended reported speech should draw on one or more of these meaning components. That is, over the course of grammaticalisation some of these semantic features become irrelevant or develop a broader interpretation.

For each of the extended meanings illustrated in **Section 2** we may indeed hypothesise that this is the case: 'demonstratedness' may serve as a source meaning for (grammatical) functions relating to prominence (cf. discourse functions), comparativity (cf. similitive) and unithood (cf. complement clause marking). Interestingly, these meanings are very close to the kinds of meaning extensions of fictive interaction which Jarque (2016), 175–181 finds in sign languages²⁴. Under our approach to evidentiality this semantic component of reported speech could extend to other functions that introduce a contrast between two events, such as temporal meanings (cf. Zeman, 2019), as well as evidential extensions themselves. The modal meaning of reported speech may further account for the multitude of attitudinal meaning extensions.

Table 1 summarises the hypotheses briefly stated above. Specifically, for the meanings listed in (19) the table suggests to which meaning components (or combinations thereof) they may be related; the Evid(ential) meaning (34b), Mod(al)meaning (34c) or Dem(onstrated) status (34a). For meaning extensions for which the respective component appears to have been completely backgrounded, the label is struck out in **Table 1**. If an Evidential, Modal or Demonstrated meaning could be interpreted as having served as input for the specific meaning extension, that is it may explain part of the extended meaning but does not fully correspond to the extended meaning itself (as in the hypotheses posited in the preceding paragraph), it has been italicised and underlined. We will not discuss these possible grammaticalisation paths in further detail; our main aim in proposing them is to suggest that despite that great variety of subtypes of extended reported speech, they may be given explanations based on a limited number of variables: the semantics of reported speech constructions and a combination of three processes, viz. semantic bleaching, recasting and rescaling.

The processes of recasting and rescaling were introduced in (33) and roughly correspond to those types of extended reported speech in which the reported speaker appears to have acquired a non-locutionary role, for example that of a 'thinker' or 'intention holder', and those in which the Report is not interpreted as a reported utterance. These two processes obviously mutually imply each other, but could still be seen as distinct diachronic pathways. **Table 1** suggests the relevance of these processes for each of the subtypes of extended reported speech.

A full analysis of the structural diachronic changes and dynamic variation in the sample languages lies beyond the scope of this

²³This property explains why elements that in other grammatical contexts do not carry any specific attitudinal properties, such as pronouns or tenses, can gain modal meanings in the context of reported speech (cf. Zemp, 2020).

²⁴While Jarque (2016) discusses 'fictive questions', not extended reported speech, the grammaticalised meanings of fictive questions correspond quite closely to the ones we attribute to the 'demonstrated' status of reported speech. We thank a reviewer for pointing out this connection.

article (but see Casartelli, 2019), but **Table 1** suggests that rescaling takes several forms: the Report unit is typically an illocutionary unit, but may be reinterpreted as a clausal unit of various kinds (in the case of, e.g., ‘highlighting’ and ‘topic’). In, for example, modal extended reported speech, R is interpreted as a proposition²⁵, and even a smaller scope unit for, for example, the aspectual subtype, which we have labelled ‘event’ in **Table 1**. In addition, the subject of the Matrix unit, typically the reported speaker, may be recast as, for example, a thinker, an intention holder, or ‘aspectual viewpoint taker’, often in combination with a rescaled R.

With this brief semantic discussion we have aimed to show that rather than constituting a scattered range of unrelated meanings, the functions attested in extended reported speech can be captured using a rather restricted set of variables that are directly related to the semantics of reported speech.

4.2 Extended Reported Speech and the Study of Perspective

As indicated in our case studies in **Section 3** and as suggested by several observations in **Section 2**, meaning extensions in reported speech often seem restricted to specific structural contexts. For example, Reesink (1993) notes with respect to Usan extended reported speech:

‘It is clear that all seven [extended] “functions” exhibit only one form of the verb ‘to say’, the medial [Same Subject] form [...] I would suggest, then, that Usan has only two functions for *qamb* ‘to say’. The first is the general function to refer to the act of speaking or telling. This allows all possible forms of the verb paradigm. The second function is what we could call a grammaticalized one, which allows only the medial Same Subject form *qamb*. This one covers all instances that refer to “inner speech” (Reesink, 1993, 223).

The relative flexibility of the ‘regular’ reported speech construction compared to reported speech structures with extended meanings in Usan is mirrored by multiple accounts. Also, decreased variation in the choice of indexical values of pronouns may covary with the extended meanings of reported speech more generally. This is the case in the example of Sanzi reported thought in (14a), which shows conflicting referential values (in itself a more common property in Caucasian languages). In (14a), while the bound pronouns in the Report have a third person referent, the free pronouns have a first person value, yet both index the same referent, viz. the person uttering the example at the current speech moment (Forker, 2019). In Sanzi this appears to be a strategy to identify specific referents both in reported speech and extended interpretations, but in the Daghestanian language Tabasaran such referential conflicts between bound pronouns and pronominal clitics appear to be restricted to reported speech, and not allowed

in (otherwise similar) forms of reported thought (Yaroshevich, 2020).

As Nikitina (2020) discovers, logophoric pronouns, which typically signal coreferentiality between a referent of the Matrix (often the subject) and the subject of the Report, are also required for extended meanings such as the inchoative interpretation in Wan (2). As we found in **Section 3.2.2**, the observation by Rumsey (1990) that the WANT interpretation of reported speech in Ngarinyin is restricted to Reports with first person subjects, a finding replicated in other Australian languages (McGregor, 2007, 2014) and elsewhere (cf. Everett, 2008, 389), also occurred in our cross-linguistic sample.

Chappell (2012), 81 explicitly proposes the following constructional frames in Sinitic which correspond to specific subtypes of extended reported speech:

- (35)
- a. Complementizer: Verb₁ Verb_{SAY}
 - b. Topic marker: SAY + Clause
 - c. (Composite) conjunctions: Conjunction - SAY
 - d. Hearsay evidentiality: Verb_{SAY} Verb₂
 - e. Irrealis marker: (if) SAY + Clause
 - f. Clause-final discourse marker: Clause + SAY

The construction frames in (35) are distinguished by word order (i.e., the position of SAY) and the specific combination of elements. An interesting example of such a combination is the conditional embedding ‘if SAY’ in (35e), which results in an irrealis reading.

It remains to be seen to what extent the subtypes of extended reported speech correlate with consistent, cross-linguistically recurring structural features. What these observations do suggest, however, is that in the languages surveyed in this paper, a number of structural elements, like those summarised in (36), can be recruited to signal a range of extended meanings.

- (36)
- a. Pronouns and other indexical elements (including tense and mood)
 - b. Integrity of Matrix and Report units
 - c. Order of Matrix and Report units
 - d. Modification
 - i. of Matrix
 - ii. of Report

These strategies are by no means a complete list of possible structural prompts for meaning extensions (e.g., prosodic distinctions are likely to occur more widely as well; also cf. Spronck, 2016), but they hold an important implication: each of the properties in (36) is associated with other aspects of the classification of reported speech constructions. For example, the indexical properties of reported speech are commonly associated with the opposition between direct speech and indirect speech (as in 37 and 38, respectively). The integration of the Matrix and Report units corresponds to a distinction between having two syntactically separate (or loosely connected) clauses as in direct speech, two more integrated clauses, as in the complementation structure of (English) indirect speech and, e.g., even further structurally integrated expressions of Matrix units, as in adverbial (or morphological) expressions of reportative evidentiality (as in 39). Finally, we have also observed that over the course of grammaticalisation, Matrix clauses may become less clearly marked, a distinction commonly associated with the opposition between types of reported speech with a

²⁵Following Boye (2012), 204 we also classify the difference between quotative and reportative evidentiality in terms of the type of embedded unit: a locution vs a proposition, respectively.

TABLE 1 | Suggested processes of semantic bleaching, rescaling and recasting in extended reported speech. The table lists for each of the subtypes of extended reported speech, which of the three semantic components of reported speech, viz. evid(entiality), mod(ality) and dem(onstratedness), are bleached, indicated by being ~~struck-out~~ or extended, in which case the relevant semantic component is *underlined and italicised*. For rescaling 'R >' indicates the type of semantic unit into which the Report is reanalysed (the precise labels 'name', 'event' etc. are indicative and should be more narrowly defined in future research). The roles indicated after 'recasting' suggest the semantic interpretation of the referent who is marked as the reported speaker in the extended reported speech construction.

| MEANING CLASS | SUBTYPE | SUGGESTED PROCESSES |
|------------------------------|---------------------|--|
| <i>naming</i> | 'Call' | Bleaching: Evid/Mod/Dem; rescaling: R > name (noun phrase); recasting: M _S > generic 'caller' |
| <i>thought</i> | 'Think' | Bleaching: Evid/Mod/Dem; rescaling: R > thought (locution/inner utterance); recasting: M _S > 'thinker' |
| <i>thought</i> | 'Want' | Bleaching: Evid/Mod/Dem; rescaling: R > wish/intention (proposition); recasting: M _S > 'intention holder' |
| <i>attitude</i> | 'Modality' | Bleaching: Evid/Mod/Dem; rescaling: R > proposition; recasting: M _S > <i>various</i> |
| <i>attitude</i> | 'irrealis' | Bleaching: Evid/Mod/Dem; rescaling: R > proposition; recasting: <i>various</i> |
| <i>evidentiality</i> | 'Quotative' | Bleaching: Evid/Mod/Dem; rescaling: R > -; recasting: |
| <i>evidentiality</i> | 'Reportative' | Bleaching: Evid/Mod/Dem; rescaling: R > proposition; recasting: |
| <i>time</i> | 'Inchoative aspect' | Bleaching: Evid/Mod/Dem; rescaling: R > proposition; recasting: 'Temporal viewpoint taker' |
| <i>time</i> | 'future tense' | Bleaching: Evid/Mod/Dem; rescaling: R > event; recasting: 'Aspectual viewpoint taker' |
| <i>valency changing</i> | 'Causation' | Bleaching: Evid/Mod/Dem; rescaling: R > event; recasting: Referent (causer) |
| <i>valency changing</i> | 'Reason' | Bleaching: Evid/Mod/Dem; rescaling: R > event; recasting: Referent (causer) |
| <i>valency changing</i> | 'Purpose' | Bleaching: Evid/Mod/Dem; rescaling: R > name; recasting: Referent (intention holder) |
| <i>clause linking</i> | 'Complementiser' | Bleaching: Evid/Mod/Dem; rescaling: R > clause; recasting: |
| <i>clause linking</i> | 'Connective' | Bleaching: Evid/Mod/Dem; rescaling: R > clause; recasting: |
| <i>information structure</i> | 'Topic' | Bleaching: Evid/Mod/Dem; rescaling: R > (sub)clause; recasting: |
| <i>information structure</i> | 'Highlighting' | Bleaching: Evid/Mod/Dem; rescaling: R > (sub)clause; recasting: |

clearly indicated source and types in which this is not the case, as in free indirect speech (as in 40), where only the Report unit is explicitly expressed.

(37) Sasha said: "I will go."

(38) They said that they would attend.

(39) Allegedly, you were there.

(40) He looked at her. Would she really leave? Her gaze froze.

Extended reported speech intersects the four types of reported speech illustrated in (37–40), but also defies this classification, with some examples not clearly belonging to any of these four classes. For the study of perspective this has the implication that in extended reported speech we see non-perspective expressions emerge, both semantically (Gentens et al., 2019, 159) and structurally, out of perspectival constructions. Reported speech typically signals that the content of the Report is grounded in a perspective other than that of the current speaker at the speech moment. For most examples of extended reported speech the perspective associated with the Matrix and the Report is the same for both unit, however, that is, that of the current speaker. Where the construction involved is still structurally clearly identifiable as reported speech it constitutes a form-function mismatch in which the typical meaning of this construction would indicate a change in perspective, but its interpretation is 'perspective persistent' in terms of Gentens et al. (2019) and Spronck et al. (2020). The loss of perspective meaning may also be iconically signalled in the linguistic structure through the various marking variations found in extended reported speech²⁶.

²⁶Note that this phenomenon complements a reverse diachronic direction that elements *within* reported speech can display with respect to perspectival interpretations: word classes and categories that do not necessarily signal perspective meanings may gain such a meaning in the context of reported speech. A particularly prominent example of such a development is formed by pronouns, which may develop evidential meanings (cf. Zemp, 2020) or take on referential meanings specific to the reported speech context (cf. Nikitina, 2012).

The examples illustrated in this study appear to suggest that extended reported speech often also operates in the categorial twilight area between direct speech and non-direct speech. Even though most authors in our survey in Section 2 consider reported speech expressions other than direct speech marked or even exceptional in the respective language, very few of the examples of extended reported speech cited are common direct speech structures. Pascual (2014), 49 makes a similar observation about her data sample: 'On the one hand, the cases discussed in this section share all the formal characteristics of direct speech. On the other hand, their possible appearance after complementizer 'that', their multifunctionality, and their type rather than token interpretation constitute features traditionally associated with indirect speech'. We would add that also *structurally*, extended reported speech often displays 'indirect-like' features.

4.3 A Speculative Story: Reported Speech as the Origin of Grammar

After having noted that extended reported speech constitutes a wide range of subtypes, that are nevertheless quite regular and can be related to a common semantic origin and (more impressionistically) share certain structural features, we would like to return to the research programme that we started out with at the beginning of this paper: the study of fictive interaction. The implication that extended reported speech has for this research programme is admittedly speculative, but to us it also seems to be the most exciting one: in extended reported speech a connection appears to emerge between the representation of other people's utterances and grammar. This allows us to propose a fundamental hypothesis about how these grammatical meanings may ultimately have arisen in the evolution of language.

Pascual (2014) convincingly demonstrates that metaphors of conversation are a frequent strategy for speakers to explain complex concepts and may affect language at any grammatical level. Furthermore, our ability to reason, according to Mercier

and Sperber (2017), arose out of a discursive need to evaluate the effectiveness of our arguments in conversation. In human evolution, this ability must have been predated by the capacity for being able to recognise the world view and knowledge of others as different from our own, popularly referred to as ‘theory of mind’ (Tomasello, 2014). Like most evolved capacities, this is not a uniquely human trait (de Waal, 2016), but it is a necessary step for the use of symbolic communication (Dor, 2017).

Built on these cognitive foundations, the assumption that language started out as situation-specific calls, developing into non-situation specific symbolic conventions for communication of ever increasing complexity (cf. Dor, 2015, ch. 8) seems relatively uncontroversial. But this scenario also assigns a central role to linguistic reflexivity in language evolution: it requires speakers to reflect on the form and meaning of what-is-said, the ability to ‘turn language on itself’ (Lucy, 1993). The type of linguistic structures specifically dedicated to this task are reported speech. If linguistic reflexivity, that is, thinking and talking about language, is at the heart of the complexification of grammar, reported speech is at the heart of language evolution, which would at once explain its universality in the languages of the world and its relation to grammatical categories, as indicated by the range of functions summarised in **Section 2**.

We do not wish to suggest that any of the languages cited in this paper represent an evolutionary early stage of grammatical development. Given the importance of metaphors of conversation in language (Pascual, 2014), grammaticalisation and semantic extension of reported speech structures may be cyclical or run parallel to other diachronic developments. We also do not suggest that in deep history all markers of, e.g., aspect or causation must have emerged out of reported speech. Rather, we would propose that the semantic components of reported speech provide a model for the lexical and grammatical meanings listed in (19). Once the communicative utility of this meaning is adopted by the speech community, it may have been marked through a special form of a reported speech construction, or a newly emerged form dedicated to this specific meaning. In this scenario, reported speech constructions may either have acted as a formal source for grammatical categories associated with the functions in (19) or a semantic model.

In order to test this hypothesis we need to further examine the semantic commonalities between reported speech and the respective grammatical categories involved in the extensions, as well as the semantic oppositions that exist between extended reported speech and morphological categories in the languages that both have, e.g., tense meanings based on reported speech forms and a separate morphological tense form.

Nonetheless, the regularity of the large range of semantic extensions of reported speech, as well as their apparent similarity to the meanings of some of the most basic grammatical categories in the languages of the world, is unlikely to be coincidental. Although the evolutionary story sketched here is inevitably speculative, we believe that it is also a plausible story about the development of grammatical complexity and constitution of grammatical categories. Above all, it motivates the importance of gaining a deeper understanding of the diversity of structures and meanings associated with extended reported speech and their relation to perspective expressions and grammar more generally.

5 CONCLUSION

In this article we have aimed to develop a typological approach to extended reported speech, highlighting both the wide-ranging forms and functions of the phenomenon and its apparent regularity. Ultimately, this leads us to suggest that extended reported speech constitutes a fertile birth environment for core grammatical meanings: the list of subtypes summarised in (19) includes lexical extensions alongside some of the most common verbal categories found in the languages of the world: evidentiality, modality, aspect/ tense, valency change, among others.

Much work remains to be done in order to gain a fuller picture of both the semantic patterns found in extended reported speech around the world, and of the structural patterns employed to express these meanings. These typological questions should be answered in dialogue with theoretical discussions about how quotation shifts perspective and what the semantic status is of the content of a Report; as well as what aspects of reported speech are conventional and which are pragmatic.

This may ultimately lead us to an understanding of why grammar is the way it is.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

Both authors discussed the content and structure of the article. DC constructed the typological sample, produced the map and collected the examples cited in **Section 4**. SS designed the study and wrote the article.

ACKNOWLEDGMENTS

We are grateful to the reviewers for their helpful and constructive comments, to Sonja Zeman for exceptional editorial support and comments, to Angela Terrill for comments and copy-editing, as well as to the participants of our bi-weekly online data workshop on reported speech (October 2020 - June 2021), particularly (in alphabetical order), Anna Bugaeva, Sonja Gipper, Guillaume Guitang, Izabela Jordanoska, Tatiana Nikitina, Rebecca Paterson, Lacina Salué, Elena Skribnik, Adam Tallman and Denys Teptiuk for comments on an earlier draft of this paper and stimulating discussion. We thank the University of Helsinki for financial support through the project ‘Language emerging from human sociality: the case of speech representation.’

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2021.624486/full#supplementary-material>

REFERENCES

- Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford etc: Oxford University Press.
- Bakhtin, M. (1981). *The Dialogic Imagination: Four Essays by M. M. Bakhtin*. Austin: University of Texas. Caryl Emerson and Michael Holquist.
- Baranova, V. (2015). Grammaticalization Paths of the Verb Gi- 'say' in Kalmyk. *Mongolica Pragmatisia. Linguistics, Ethnolinguistics, Religion Cult.* 8 (2), 57–76.
- Bickel, B., Comrie, B., and Haspelmath, M. (2008). The Leipzig Glossing Rules: Conventions for Interlinear Morpheme-By-Morpheme Glosses. Most recent version available at: <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>.
- Birchall, J. (2018). Historical Change in Reported Speech Constructions in the Chapacuran Family. *Jhl* 8 (1), 7–30. doi:10.1075/jhl.00003.bir
- Bon, N. (2014). *Une grammaire de la langue stieng, langue en danger du Cambodge et du Vietnam*. Ph. D. thesis. Université Lumière Lyon 2.
- Boye, K. (2012). *Epistemic Meaning: A Crosslinguistic and Functional-Cognitive Study*. Berlin/Boston: De Gruyter Mouton.
- Bugaeva, A. (2008). Reported Discourse and Logophoricity in Southern Hokkaido Dialects of Ainu. *Gengo Kenkyu* 133, 31–75.
- Casartelli, D. (2019). *Beyond Speech Reporting: A Crosslinguistic Analysis of Non-locutionary Meanings in Reported Speech*. Ph. D. thesis, Helsinki: University of Helsinki In preparation.
- Chappell, H. (2012). "Say in Sinitic: From *Verba Dicendi* to Attitudinal Discourse Markers," in *Grammaticalization and (Inter-)subjectification*. Editors J. van der Auwera and J. Nuyts (Wetteren: Universa Press), 81–110.
- Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford etc: Oxford University Press.
- Cinque, G. (2013). Cognition, Universal Grammar, and Typological Generalizations. *Lingua* 130, 50–65. doi:10.1016/j.lingua.2012.10.007
- Clark, H. H. (2016). Depicting as a Method of Communication. *Psychol. Rev.* 123 (3), 324–347. doi:10.1037/rev0000026
- Clark, H. H., and Gerrig, R. J. (1990). Quotations as Demonstrations. *Language* 66, 764–805. doi:10.2307/414729
- Clendona, M. (2014). *Worrorra: A Language of the north-west Kimberley Coast*. Adelaide: University of Adelaide Press.
- Coate, H. H. J. (1966). The Rai and the Third Eye North-west Australian Beliefs. *Oceania* 37 (2), 93–123. doi:10.1002/j.1834-4461.1966.tb01790.x
- Cohen, D., Simeone-Senelle, M.-C., and Vanhove, M. (2002). "11. The Grammaticalization of 'say' and 'do,'" in *Reported Discourse. A Meeting Ground for Different Linguistic Domains*. Editors T. Güldeman and M. von Stechow (Amsterdam/Philadelphia: John Benjamins), 227–251. doi:10.1075/tsl.52.17.coh
- Cysouw, M., and Good, J. (2013). Languoid, Doculect, and Glossonym: Formalizing the Notion 'language'. *Lang. Documentation Conservation* 7, 331–360.
- Davidson, K. (2015). Quotation, Demonstration, and Iconicity. *Linguist Philos.* 38, 477–520. doi:10.1007/s10988-015-9180-1
- De Smet, H. (2012). The Course of Actualization. *Language* 88 (3), 601–633. doi:10.1353/lan.2012.0056
- de Waal, F. (2016). *Are We Smart Enough to Know How Smart Animals Are?* New York/London: W. W. Norton.
- Deibler, E. (1971). Uses of the Verb 'to Say' in Gahuku. *Kivung* 4 (2), 101–110.
- Dhakai, D. N. (2012). *Darai Grammar, Volume 489 of Languages of the World/Materials*. München: Lincom.
- Dik, S. C. (1997a). *The Theory of Functional Grammar: Part 1 the Structure of the Clause*. Berlin/New York: De Gruyter Mouton.
- Dik, S. C. (1997b). *The Theory of Functional Grammar: Part 2 Complex and Derived Constructions*. Berlin/New York: De Gruyter Mouton.
- Dor, D. (2015). *The Instruction of Imagination: Language as a Social Communication Technology*. Oxford etc: Oxford University Press.
- Dor, D. (2017). The Role of the Lie in the Evolution of Human Language. *Lang. Sci.* 63, 44–59. doi:10.1016/j.langsci.2017.01.001
- Dryer, M. S. (1989). Large Linguistic Areas and Language Sampling. *Stud. Lang.* 13 (2), 257–292. doi:10.1075/sl.13.2.03dry
- Dryer, M. S., and Haspelmath, M. (2013). in *The World Atlas of Language Structures Online* (Leipzig: Max Planck Institute for Evolutionary Anthropology).
- Ershova, K. (2012). Reported Speech and Reportative Grammaticalization in Besleney Kabardian. In Proceedings of the Second Central European Conference in Linguistics for Postgraduate Students, 71–87.
- Everett, D. L. (2008). "Wari' Intentional State Constructions," in *Investigations of the Syntax-Semantics-Pragmatics Interface*. Editor R. D. Van Valin Jr. (Amsterdam/Philadelphia: John Benjamins), 381–409. doi:10.1075/slcs.105.27.eve
- Forker, D. (2019). Reported Speech Constructions in Sanzhi Dargwa and Their Extension to Other Areas of Grammar. *Sprachwissenschaft* 44 (2).
- Gentens, C., Sansiñena, M. S., Spronck, S., and Van Linden, A. (2019). Irregular Perspective Shifts and Perspective Persistence, Discourse-Oriented and Theoretical Approaches. *Prag.* 29 (2), 155–169. Special issue on perspective-indexing constructions: discourse and theory. doi:10.1075/prag.18050.gen
- Gentens, C. (2020). *The Factive-Reported Distinction in English*. New York: De Gruyter.
- Güldemann, T. (2008). *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Berlin: Mouton de Gruyter.
- Heine, B., and Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge etc: Cambridge University Press.
- Hengeveld, K. (1989). Layers and Operators in Functional Grammar. *J. Ling.* 25 (1), 127–157. doi:10.1017/s0022226700012123
- Hooper, J. B., and Thompson, S. A. (1973). On the Applicability of Root Transformations. *Linguistic Inq.* 4 (4), 465–497.
- Hsieh, F. (2012). On the Grammaticalization of the Kavalan SAY Verb Zin. *Oceanic Linguistics* 51 (2), 464–489. doi:10.1353/ol.2012.0025
- Jakobson, R. (1957). *Shifters, Verbal Categories and the Russian Verb*. Cambridge, MA: Department of Slavic Languages and Literatures Harvard University.
- J. A. Lucy (1993). in *Reflexive Language: Reported Speech and Metapragmatics* (Cambridge etc: Cambridge University Press).
- Jarque, M. J., and Pascual, E. (2015). Conversational Structures Marking Evidential Values in Catalan Sign Language. *eHumanista/IVITRA* 8, 421–445.
- Jarque Moyano, M. J. (2016). "What About? Fictive Question-Answer Pairs for Non-Information-Seeking Functions Across Signed Languages," in *The Conversation Frame: Forms and Functions of Fictive Interaction*. Editors Pascual and Sandler. (Amsterdam/Philadelphia: John Benjamins), 171–192. doi:10.1075/hcp.55.09.jar
- Klamer, M. (1998). *A Grammar of Kampera*. Berlin, New York: Berlin: Mouton Grammar Library/Mouton de Gruyter.
- Klamer, M. (2000). How Report Verbs Become Quote Markers and Complementisers. *Lingua* 110, 69–98. doi:10.1016/s0024-3841(99)00032-7
- Konnerth, L. (2020). Recycling through Perspective Persistence in Monsang (Trans-himalayan). *FoL* 27 (1), 55–77. doi:10.1075/foL20002.kon
- Kuteva, T., Heine, B., Hong, B., Long, H., Narrog, H., and Rhee, S. (2019). *World Lexicon of Grammaticalization*. Cambridge University Press.
- Larson, M. L. (1978). *The Functions of Reported Speech in Discourse*. Dallas: Summer Institute of Linguistics.
- Lefebvre, C., and Brousseau, A.-M. (2002). *A Grammar of Fongbe of Mouton Grammar Library*. Berlin, New York: Berlin: Mouton de Gruyter.
- Loughnane, R. (2009). *A Grammar of Oksapmin*. Ph. D. thesis, Melbourne: University of Melbourne.
- Martin Haspelmath, M. (2010). Comparative Concepts and Descriptive Categories in Crosslinguistic Studies. *Language* 86 (3), 663–687. doi:10.1353/lan.2010.0021
- Matić, D., and Pakendorf, B. (2013). Non-canonical SAY in Siberia. *Stud. Lang.* 37 (2), 356–412.
- McGregor, W. B. (2007). "2. A Desiderative Complement Construction in Warrwa," in *Language Description, History and Development*. Editors J. Siegel, J. Lynch, and D. Eades (Amsterdam/Philadelphia: John Benjamins), 27–40. doi:10.1075/cll.30.08.mcg
- McGregor, W. B. (2019). Reported Speech as a Dedicated Grammatical Domain - and Why Defenestration Should Not Be Thrown Out the Window. *Linguistic Typology* 23 (1), 207–219. doi:10.1515/lingty-2019-0010
- McGregor, W. B. (2014). "The 'Say, Do' Verb in Nyulnyul, Warrwa, and Other Nyulnyulan Languages Is Monosemic," in *Events, Arguments and Aspects: Topics in the Semantics of Verbs*. Editor K. Robering (Amsterdam/Philadelphia: John Benjamins), 301–328. doi:10.1075/slcs.152.08gre
- Mercier, H., and Sperber, D. (2017). *The Enigma of Reason*. Cambridge, MA: Harvard University Press.

- Miestamo, M., Bakker, D., and Arppe, A. (2016). Sampling for Variety. *Linguistic Typology* 20 (2). doi:10.1515/lingty-2016-0006
- Moroz, G. (2017). *Lingtypology: Linguistic Typology and Mapping*.
- Nikitina, T. (2020). Logophoricity and Shifts of Perspective. *FoL* 27 (1), 78–99. doi:10.1075/foL.20001.nik
- Nikitina, T. (2012). Personal Deixis and Reported Discourse: Towards a Typology of Person Alignment. *Linguistic Typology* 16 (2), 233–263.
- Pascual, E. (2007). “Fictive Interaction Blends in Everyday Life and Courtroom Settings,” in *Mental Spaces in Discourse and Interaction*. Editors T. Oakley and A. Hougaard (Amsterdam/Philadelphia: John Benjamins), 79–107.
- Pascual, E. (2014). *Fictive Interaction: The Conversation Frame in Thought, Language, and Discourse*. Amsterdam/Philadelphia: John Benjamins.
- Pascual, E. (2002). *Imaginary Dialogues: Conceptual Blending and Fictive Interaction in Criminal Courts*. Utrecht: LOT.
- Pascual, E., and Sandler, S. (2016a). In *The Conversation Frame: Forms and Functions of Fictive Interaction* (Amsterdam/Philadelphia: John Benjamins).
- Pascual, E., and Sandler, S. (2016b). “Fictive Interaction and the Conversation Frame,” in *The Conversation Frame: Forms and Functions of Fictive Interaction*. Editors E. Pascual and S. Sandler (Amsterdam/Philadelphia: John Benjamins), 3–22. doi:10.1075/hcp.55.01pas
- Pawley, A., and Bulmer, R. (2011). *Research School of Pacific and Asian Studies*. Canberra: Australian National University. A Dictionary of Kalam With Ethnographic Notes, Volume 630 of *Pacific Linguistics*.
- Pennington, R. (2018). *A Grammar of Ma Manda: A Papuan Language of Morobe Province, Papua New Guinea, Volume 18 of Outstanding Grammars from Australia*. München: LINCOM.
- Reesink, G. P. (1993). ‘Inner Speech’ in Papuan Languages. *Lang. Linguistics Melanesia* 24, 217–225.
- Rijkhoff, J., Bakker, D., Hengeveld, K., and Kahrel, P. (1993). A Method of Language Sampling. *Sl* 17 (1), 169–203. doi:10.1075/sl.17.1.07rij
- Rumsey, A. (1982). *An Intra-Sentence Grammar of Ungarinjin*, North-Western Australia. Canberra: Pacific Linguistics 92 (2), 157–166. doi:10.1525/aa.1990.92.2.02a00060
- Rumsey, A. (1990). Wording, Meaning, and Linguistic Ideology. *Am. Anthropologist* 92 (2), 346–361. doi:10.1525/aa.1990.92.2.02a00060
- Sandler, S., and Pascual, E. (2019). In the Beginning There Was Conversation: Fictive Direct Speech in the Hebrew Bible. *Prag.* 29 (2), 250–276. doi:10.1075/prag.18047.san
- Saxena, A. (1988). On Syntactic Convergence: The Case of the Verb ‘say’ in Tibeto-Burman. In Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society, 375–388. doi:10.3765/bls.v14i0.1781
- Spronck, S. (2015). *Reported Speech in Ungarinjin: Grammar and Social Cognition in a Language of the Kimberley Region*. Western Australia: Ph. D. thesis The Australian National University.
- Spronck, S. (2016). “Evidential Fictive Interaction (In Ungarinjin and Russian),” in *The Conversation Frame: Forms and Functions of Fictive Interaction*. Editors E. Pascual and S. Sandler (Amsterdam/Philadelphia: John Benjamins), 255–275. doi:10.1075/hcp.55.13spr
- Spronck, S., and Nikitina, T. (2019). Reported Speech Forms a Dedicated Syntactic Domain. *Linguistic Typology* 23 (1), 119–159. doi:10.1515/lingty-2019-0005
- Spronck, S., Van Linden, A., Gentens, C., and Sansiñena, M. S. (2020). Perspective Persistence and Irregular Perspective Shift. *FoL* 27 (1), 1–6. doi:10.1075/foL.20005.spr
- Thornes, T. J. (2003). *A Northern Paiute Grammar and Texts*. Eugene: Ann Arbor: Ph. D. thesis University of Oregon.
- Tomasello, M. (2014). *A Natural History of Human Thinking*. Cambridge/London: Harvard University Press.
- van der Voort, H. (2002). “The Quotative Construction in Kwaza and its (De-) Grammaticalisation,” in *Current Studies on South American Languages Indigenous Languages of Latin America*. Editors M. Crevels, S. van de Kerke, S. Meira, and H. van der Voort, 307–328.
- Völlmin, S. (2017). *Towards a Grammar of Gumer: Phonology and Morphology of a Western Gurage Variety*. Ph. D. thesis, Zürich: University of Zürich.
- Vydrina, A. (2017). *A Corpus-Based Description of Kakabe, a Western Mande Language: Prosody in Grammar*. Paris: Ph. D. thesis INALCO.
- Vygotsky, L. S. (1987). *Thinking and Speech. The Collected Works of Lev Vygotsky*. New York, NY: Plenum Press.
- Wiemer, B. (2018). *Catching the Elusive: Lexical Evidentiality Markers in Slavic Languages (A Questionnaire Study and its Background)*. Berlin: Peter Lang.
- Wilson, J. (2017). *A Grammar of Yeri: A Torricelli Language of Papua New Guinea*. Ph. D. thesis, Buffalo: State University of New York at Buffalo.
- Yaroshevich, S. (2020). Is Thinking like Saying? the Case of Tabasaran. Paper read at the 53rd Annual Meeting of the *Societas Linguistica Europaea* in the workshop ‘The grammar of thinking: Comparing reported thought and reported speech across languages’.
- Zeman, S. (2019). The Emergence of Viewpoints in Multiple Perspective Constructions. *Prag.* 29 (2), 226–249. doi:10.1075/prag.18051.zem
- Zemp, M. (2020). Evidentials and Their Pivot in Tibetic and Neighboring Himalayan Languages. *FoL* 27 (1), 29–54. doi:10.1075/foL.20003.zem

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Spronck and Casartelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GLOSSARY

{} elements in the ME gloss that are not part of the extended reported speech construction

- ADD** additive (Besleney Kabardian)
- AF** agent focus (Kavalan)
- ANAPH** anaphoric pronoun
- ART** article (Kakabe)
- CM** conjugation marker (Warrwa)
- CONJV** conjunctive (Stieng)
- CVB.MOD** modal converb or converb of manner (Kalmyk)
- DIR** directional (Besleney Kabardian, Mandarin)
- FP** far past (Usan)
- HOR/HORT** hortative (Darai, Kalam)
- ICVB**
imperfective converb (Sanzhi)
- IF** immediate future (Oksapmin)
- IP** instrumental prefix (Northern Paiute)
- LF** locative focus (Kavalan)
- LOG** logophoric pronoun
- 'M'** verbal '-m' suffix (Gumer) (Völlmin, 2017, 152)
- MM** middle marker (reflexive, reciprocal, passive, middle; Northern Paiute)
- MIN** minimal number (Warrwa)
- MIR** mirative (Darai)
- NF** non-finite (Newari)
- NFUT** non-future
- NONVIS.EV** nonvisual evidential (Ainu)
- NPC** non-past conjunct (Newari)
- OBL** oblique stem marker (Sanzhi)
- PART** particle (Darai)
- PD** past disjunct (Newari)
- PN** personal name (Ma Manda)
- POT** potential mood
- PRET** preterite (Sanzhi)
- PRIOR** priorite (Kalam)
- PRT** particle (Wan, Mandarin)
- RECG** recognitional (Oksapmin)
- RED** reduplication (Usan)
- SEQ** sequential (Darai)
- SS** same subject (Usan)
- UF** uncertain future (Usan)
- W-CLASS** second ("Wu-") neuter gender class (Ngarinyin)