

# Gabrielle Hodge\* and Kearsy Cormier

## Reported speech as enactment

<https://doi.org/10.1515/lingty-2019-0008>

## 1 Introduction

We examine the proposal by Spronck and Nikitina (this issue) that the phenomenon known as ‘reported speech’ constitutes a dedicated syntactic domain. This domain is defined as a “dedicated syntactic relation that differs from other sentential syntactic structures”, containing elements which are “primarily conventionalized and conditioned by grammar in a narrower sense”. Their stated goal is to “identify and classify phenomena occurring in the context of reported speech, and to propose benchmarks for establishing reported speech as a cross-linguistic category”. We broadly agree with the definition of reported speech proposed as useful for investigating how specific languages constrain the inference of this particular aspect of meaning via ostensive and inferential communicative acts (LaPolla 2003). The proposed definition enables functional identification of all kinds of reported speech practices while making no assumptions about how these practices formally manifest within different language ecologies. However, we disagree with the case for syntactic exceptionalism, i.e. the category ‘reported speech’ as a dedicated syntactic domain. Our main concern is that this claim downplays important evidence regarding the unified and multimodal ‘semiotic repertoire’ (Kendon 2014; Kusters et al. 2017) available for reporting utterances, thoughts, feelings, attitudes and actions across diverse languages – including deaf signed languages – of which the highly grammaticalised and conventional encoding of reported speech utterances is just one part.

We argue that acts of reported speech are primarily *depictions*, which may not necessarily involve conventionally symbolic form-to-meaning mappings. Indeed, they are often entirely non-conventional bodily actions that are heavily dependent on the unfolding context for interpretation: “[t]hey use material gradiently so that certain changes in form imply analogical differences in meaning...to interpret depictions, we imagine what it is like to see the thing depicted” (Dingemanse 2015: 950). These depictions may (or may not) be framed by more conventional

---

\*Corresponding author: **Gabrielle Hodge** [*gæbɪl'el 'hɒdʒ*], Deafness Cognition and Language Research Centre, University College London, 49 Gordon Square, London, WC1H 0PD, UK, E-mail: [g.hodge@ucl.ac.uk](mailto:g.hodge@ucl.ac.uk)

**Kearsy Cormier** [*'ki:si 'kɔ:mi:ɪ*], Deafness Cognition and Language Research Centre, University College London, 49 Gordon Square, London, WC1H 0PD, UK, E-mail: [k.cormier@ucl.ac.uk](mailto:k.cormier@ucl.ac.uk)

*indications* (such as indexical finger-pointing actions or deictic speech markers) and/or *descriptions* (such as conventionalised quotative markers or verbs of saying) (Clark 1996; Ferrara & Hodge 2018). Thus, while it is certainly common for speakers of various languages to signal reported speech via conventionalised morphosyntactic strategies, and for the utterances reported to contain conventionally symbolic forms, these patterns do not represent the full range of reported speech practices observed for either signers or speakers. Instead, it is the (potentially non-conventional) semiotic properties of depiction which are foundational to all acts of reported speech (not conventionalised syntactic properties). We suggest that recruiting Spronck and Nikitina's proposed definition of reported speech for a functional comparative concept (Haspelmath 2010; Croft 2016) ENACTMENT might offer a “modality-agnostic” (Dingemanse 2019; see also Okrent 2002) semiotic method for describing and comparing language-specific strategies for reporting self and others outside the current speech moment.

## 2 The case for reported speech as a dedicated syntactic domain

Spronck and Nikitina strive to unite functional and formal frameworks for describing ‘reported speech’. They consider a range of definitions, leading with an older, sociological notion developed within the Bakhtinian group active in the early twentieth century: “speech within speech, utterance within utterance, and at the same time also *speech about speech, utterance about utterance*” (italics in the original, Vološinov 1973: 115). Spronck and Nikitina therefore propose the following definition (which includes aspects of both form and function) as compatible with most general treatments in the spoken language literature:

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

However, Spronck and Nikitina acknowledge three problems with this definition: (1) it relies too much on the prototypical projection of spoken utterances in reported speech constructions, while excluding other epistemological senses such as ‘thinking’ and ‘wanting’; (2) it obscures the great diversity in the expression of reported speech that can be observed – many of which may be marked by “minimal linguistic means or even extra-linguistic means such as

eye gaze or gesture”; and (3) it does not differentiate between reported speech as a discourse act and as a grammatical construction. They offer eight cross-linguistic observations they claim differentiate the syntax of reported speech from other syntactic categories (presumably only in spoken languages, since deaf signed languages are not mentioned at all).

The authors conclude that all reported speech components minimally include: (a) a semiotic status of ‘demonstratedness’; (b) an evidentiality component (reflecting a deictic relation between the speech event reported and the original event); and (c) a modality component (reflecting the speaker’s epistemic evaluation of the reported utterance). We agree this definition is sufficiently narrow to capture the range of reported speech strategies documented by Spronck and Nikitina, and general enough to be effectively operationalised across data from both signed and spoken languages (even though they only address the latter). However, it also captures much more than just highly grammaticalised and conventionalised morphosyntactic strategies for reporting speech, and as such cannot constitute a dedicated syntactic domain as defined by the authors – at least not without excluding important aspects of semiotic diversity for signalling description, indication and depiction across signed and spoken languages (Clark 1996; Kendon 2014; Kusters et al. 2017; Ferrara & Hodge 2018).

### 3 Reported speech as depiction signalled via multimodal repertoires of enactment

After detailing many aspects of reported speech across spoken languages, Spronck and Nikitina acknowledge that “reported speech appears to be accompanied disproportionately by gesture and prosodic cues ... and these need to be given a proper and principled place in its description”. We suggest that the strong prevalence of gesture and prosodic cues in reported speech is due to the highly improvised and mimetic nature of depiction (Clark & Gerrig 1990; Metzger 1995; Clark 2016). Spronck and Nikitina also suggest there is a dearth of research into the multimodality of reported speech, and claim that “gesture or prosody have not yet been convincingly demonstrated to contribute to its marking”. There is, however, an extensive literature describing how both signers and speakers coordinate a unified, multimodal repertoire with or without more conventionalised semiotics to gradiently depict who did what to whom and how during their face-to-face interactions, some of which we summarise briefly here.

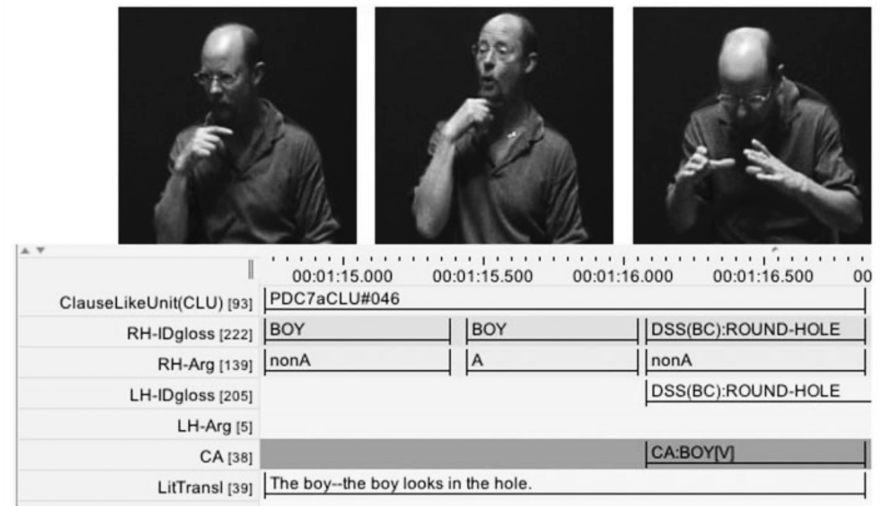
Both signers and speakers combine bodily movements, postures and eye gaze to ‘re-construct’ past or future actions and dialogue to ‘show’ characters

and events. These re-constructions may be grounded in real-life occurrences, or they may be imagined by the signer or speaker. During constructed action, one enacts a non-linguistic action (“quotes an action”). During constructed dialogue (i.e. quotative constructed action) one enacts a language event (“quotes signs or words”). Clark and Gerrig (1990: 782) give the following example: “*I got out of the car, and I just [demonstration of turning around and bumping his head on an invisible telephone pole].*” In this case, the speaker uses enactment to depict an earlier event, first framing it with a spoken English utterance, then enacting the (non-conventional) visible bodily action. In doing so, the speaker achieves a different effect compared to simply signalling primarily via conventionalised description: “*I got out of the car, and I turned around and bumped my head on a telephone pole.*” Enactment enables speakers to ‘show’ their interactants what happened, rather than just ‘tell’ it (Cormier et al. 2013).

Similarly, D’Arcy (2015: 44) provides the following example of how English speakers might construct dialogue: *I said, “Is your dad gonna sell it?” And he says “Yeah, he’s gotta go.”* (female, b. 1959). In the framed spoken English examples here, voice quality and/or visible bodily actions (not typically included in written transcriptions) are crucial for the moment-by-moment recognition of the speaker’s re-enactment of an earlier language event. A key difference between this example and the head-bumping example documented in Clark and Gerrig (1990) is that the head-bumping action constitutes an entirely non-conventional report: it is a ‘singular event’ during which interactants enchronically interpret a form as ‘standing for’ a meaning (Kockelman 2005; Enfield 2009). It does not have properties of conventionalised symbolism, i.e. meanings that are additional or predictable from the value of their form given a particular context (see also Johnston & Schembri 2010). Conversely, the examples from D’Arcy (2015) are more precisely *depictions of prior acts of description*. Yet in both cases, each depiction (via enactment) of the earlier event indexes both the original act and any subsequent depictions (Tannen 1989; Clark 1996; Ferrara & Hodge 2018).

Enactment is also used by signers – though typically through the coordination of visible bodily actions produced with the body, mouth, head and arms (i.e. not with speech) (e.g. Quinto-Pozos 2007; Cormier et al. 2013; Ferrara & Johnston 2014). Signers often rely on enactment to show and imply semantic relations between participants and events, for example, sometimes without any explicit lexical or morphosyntactic (i.e. conventionalised) encoding of these relations (Hodge & Johnston 2014). Figure 1, an example from Auslan (a deaf signed language of Australia), illustrates how signers may integrate conventionalised descriptive semiotics with co-occurring and highly improvised depictive enactment. Here a signer produces the lexical Auslan sign BOY, and then uses his hands and body to visibly enact the boy looking into a round hole in the ground,



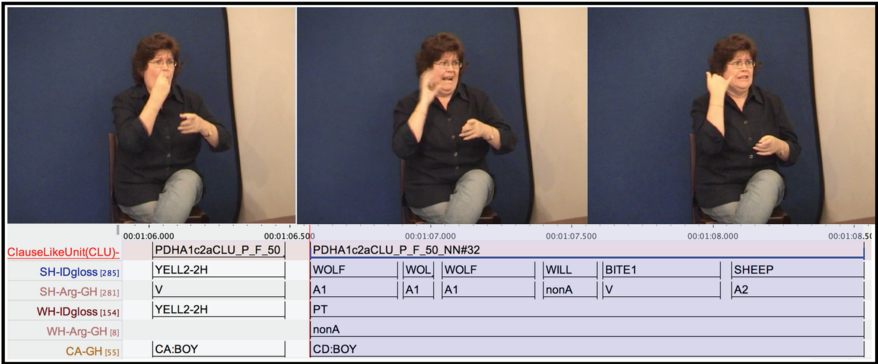


**Figure 1:** Constructed action framed via lexical sign BOY (Ferrara & Johnston 2014: 206).  
Literal translation: *boy* [demonstration of looking into a hole in the ground].

i.e. literally *boy* [demonstration of looking into a hole in the ground] (Ferrara & Johnston 2014: 206). Note the schematic similarity of this example with the earlier example from Clark and Gerrig (1990): both involve a lexical description which frames or indexes a visible bodily depiction.

Figure 2 demonstrates constructed dialogue – i.e. a depiction of a speech utterance – in Auslan (Hodge 2014: 193). Here the signer is retelling an event from the Aesop’s Fable *The boy who cried wolf*. The signer first points to a specific, meaningful location in her signing space in which the referent ‘boy’ was previously established. She then directs the lexical sign YELL to a different location in her signing space, where the referent ‘villagers’ were earlier established. Finally, the signer produces the utterance expressed by the boy to the villagers. Altogether the signer’s utterance can be translated as: *he yelled: “Wolf! Wolf! The wolf will catch the sheep!”* The enactment here co-occurs with the quotation. In both examples (Figures 1 and 2), the use of enactment enriches or even replaces meaning that may be encoded using more conventional semiotic resources, such as lexical signs or words, and it is often framed by more conventionalised aspects of language use.

These two Auslan examples also differ in how they are framed. As demonstrated in the target article, language-specific strategies for framing enactments vary widely. Some strategies may be more conventionalised and do more encoding work than others, e.g. the varied yet fully conventionalised morphosyntactic



**Figure 2:** Constructed dialogue indexed via lexical sign and English mouthing (Hodge 2014: 193). Literal translation: *(he) yelled: “Wolf! Wolf! The wolf will catch the sheep!”*

strategies used by speakers of Megeb Dargwa (examples 5–7 in the target article) and those used by the English speakers referenced in Clark and Gerrig (1990) and D’Arcy (2015) above. However, other strategies may involve less encoding work. In many signed languages, including Auslan, it is common for enactments to be framed by: (a) a finger-pointing action indexing a location in the space in front of the signer’s body, where a referent had previously been established; (b) a nominal form describing the referent (with or, as in Figure 1, without a verb of saying); or (c) a verbal predicate such as a verb of saying (e.g. Figure 2), perhaps in conjunction with other nominal forms. These strategies are similar to those used by speakers of Indonesian, Javanese and Sundanese, for example, who may frame constructed dialogue with indexical forms such as *kalau dia* (‘as for her/him’) and nominalised forms such as *katanya* (‘the word’) and *ngomongnya* (‘the speaking’) (Djenar et al. 2018: 292). In other words, it is not necessary for a fully conventional nominal or predicative element to be present for a token of enactment to be effectively indexed. Indeed, one commonality across the examples described here and in the target article is that all frames are fundamentally indexical, enabling signers and speakers to ‘point to’ the corresponding enactment.

As Spronck and Nikitina acknowledge, many instances of reported speech/constructed dialogue may also be unframed, as in the examples from speakers of Nyulnyul (example 8). The authors refer to these examples as ‘defenestrated’ (although this term problematically implies that framing is obligatory but sometimes removed), and wonder “how can the hearer still identify the correct reported speaker in the defenestrated clauses in (8)?” They conclude it is possible because Nyulnyul uses a “typologically common” strategy of re-constructing the reported exchange as question and answer pairs. In this case, “the illocutionary values of

the subsequent sentences imply distinct speakers”. But this is precisely our point – the illocutionary force of these utterances only becomes salient via the speaker’s multimodal re-enactment of an interaction between an emu and his tormentors, not by any specific feature of their syntax. The Nyulnyul speaker is depicting the interaction as they imagine it, and this must surely be reflected in their multimodal visible and/or audible expression, including gesture and prosody (which unfortunately is not transcribed). It is also common for enactments in signed languages to be unframed. In all these cases, the multimodal resources coordinated by the speaker or signer do much of the work of ensuring the demonstration is salient as a token depiction of an imagined event. These often co-occur and are closely timed with more conventional aspects of speech/sign, especially when used for direct quotation (see e.g. Sidnell 2006; Sams 2010; Stec et al. 2016).

The analyses of enactment in the examples presented here all align with the schema MATRIX [INDEX]: REPORT [ICON] abstracted from Spronck (2017) and the target article. However, note that the specific Auslan instantiations of this schema as exemplified in Figures 1 and 2 also mirror non-enacted patterns used by Auslan signers: there is nothing syntactically exceptional about them. Hodge (2014) investigated clause argument patterns in Auslan retellings and found no structural differences between clauses produced with or without enactment. By way of further illustration, consider the following two utterances attested in the Auslan Corpus (Figures 3 and 4).<sup>1</sup>



Manual sign: YELL-----  
Literal translation: (the boy) yelled (to village people)

**Figure 3:** Utterance with lexical sign YELL only (MBHA1c2a: 00:52:949; in this case, the signer is looking at his addressee).

<sup>1</sup> These examples can be viewed in the online collection of the Auslan Corpus: <https://elar.soas.ac.uk/Collection/MPI55247>



Manual sign: YELL-----  
 Constructed action: panicking-----  
 Literal translation: *(the boy panicked) "yelling-screaming-wildly" (to village people)*

**Figure 4:** Utterance with lexical sign YELL recruited into constructed action (PGMB1c2a: 01:35:920; in this case, the signer is looking away from his addressee and towards an imagined place where there is a wolf).

Figure 3 is a straightforward declarative utterance. Figure 4 is what one might call ‘quotative constructed action’. Yet the patterning of manual signs in both examples is identical: as with the unframed Nyulnyul examples, it is the visible and/or auditory enactment done by the signer or speaker which creates the ‘reporting’ effect. There is nothing in the form or order of these manual signs which ‘marks’ there being a syntactic relation of reported speech. In the case of enactments framed using a single finger-pointing action, the indexing of the enactment is done solely by pointing to a contextually-meaningful location in the signing space in front of the signer’s body, with the handshape of the pointing action being the only conventional aspect of this move. As all semiotic elements are necessary for interpretability of these multimodal utterances, it is not possible to include only the conventionalised semiotics while excluding the less conventional or even non-conventional semiotics. It is also not possible to draw a ‘boundary’ between them – all are relevant to understanding what is being communicated and all contributing forms vary gradiently (Kendon 2014). For example, the strength or intensity of the enactment may be characterised as full, reduced or subtle depending on each instantiated usage (Cormier et al. 2015).

These patterns of doing ‘reported speech’ (more precisely described as ‘constructed action’ and ‘constructed dialogue’) have also been observed for other related and unrelated deaf signed languages, e.g. British Sign Language (Cormier et al. 2013) and Norwegian Sign Language (Ferrara & Halvorsen 2017), as well as face-to-face interactions between hearing speakers of languages other than English, e.g. Spanish (Cameron 1998), Dutch (Mazeland 2006), and Korean (Park 2009). It is therefore important to note that these patterns are more likely a consequence of the quintessentially face-to-face nature of interactions involving

deaf people and signed language (Johnston 1996: 7) – and the fundamental multimodality of human communication in general (e.g. Stec et al. 2016; Goldin-Meadow & Brentari 2017) – ahead of any claim to conventionalised syntactic relations. Indeed, the availability of space during face-to-face signed language interactions has been suggested as “a fact that may influence, and even constrain, the linguistic [i.e. communicative] system in other ways” (Johnston 1996: 1). It is therefore debateable whether opportunities for reported speech acts to be constrained more conventionally (e.g. via syntactic relations) will (or will ever need to) arise for signers who are always communicating face-to-face, especially when other complex factors influencing signed language use and transmission are considered (see Schembri et al. 2018).

These explanatory factors maintain continuity with the ontogeny of human communication. Consider, for example, similarities between the Auslan example in Figure 1 with the pointing/one-word and ritualised gesture ensembles observed during adult-child interactions (Tomasello 2003). Other explanatory factors may also account for the language-specific patterns which arise from varied reported speech practices, including temporal iconicity, information structure, multimodal utterance composition, the semiotics of making meaningful use of space, diachronic and ontogenic adaptations (see e.g. Haiman 1985; LaPolla 2003; Enfield 2009; Givón 2009; and note that this list is not meant to be exhaustive). As reported speech acts are not necessarily constrained syntactically, and it cannot be assumed that this convention will manifest cross-linguistically, we cannot accept the proposal for reported speech as constituting a “dedicated syntactic domain”. We suggest instead that Spronck and Nikitina’s proposed definition of reported speech be recruited for a functional comparative concept ENACTMENT (Haspelmath 2010; Croft 2016). This would enable a modality-agnostic and more egalitarian method for identifying and understanding language-specific patterns that arise, which can then be compared typologically. This category aligns with the three minimal elements proposed by Spronck and Nikitina (i.e. a semiotic status of ‘demonstratedness’, an evidentiality component, and a modality component, none of which are defined by degree of conventionality) and can be effectively operationalised to investigate all kinds of signed and spoken language data (see Ferrara & Hodge 2018).

## 4 Conclusion

We have argued that there is no need to posit reported speech as a dedicated syntactic domain. Instead, reported speech may be observed as part of the broader phenomenon of ‘mimetic re-enactment’ (Clark 1996; D’Arcy 2015;

Stec et al. 2016). This facilitates inclusion of related but oft-overlooked phenomena (such as what might be considered ‘reported action’, as in the English and Auslan examples discussed above).

It is curious that multimodality and semiotic diversity were not highlighted more in Spronck and Nikitina’s proposal. This may be a direct consequence of working within a paradigm that posits boundaries between ‘language’ and ‘gesture’, ‘linguistic’ and ‘non-linguistic’, see e.g. their definition of semantics as “any type of meaning that can be shown to be coded through a conventional form”. This presupposes that conventionality manifests identically across language ecologies, when it is in fact a gradient property of different communicative acts. This paradigm also privileges languages (and speakers) with highly grammaticalised marking of reported speech (and other discourse functions) over languages with users who rely on more pragmatically oriented and contextually-dependent strategies to constrain various aspects of meaning (see LaPolla 2003). Description of different strategies for signalling meaning can indeed be operationalised via a “comparative semiotics” of human communication and language use (Kendon 2014) – there is no need for aprioristic ‘agreements’ about what is or is not ‘language’. This is particularly necessary for considering the semiotic repertoires available to diverse humans, including hearing non-signers and also deaf signers (see Kusters et al. 2017; Ferrara & Hodge 2018).

Once we have a better understanding of how enactment is achieved by both speakers and signers of diverse languages, we may get closer to potentially ecologically-specific and/or universal understandings of why we do it (LaPolla 2003). Why depict – via enactment – instead of describe? What power do depiction and description afford in different contexts for different people? These questions are more in line with Vološinov’s original sociological (and Marxist) motivations for understanding and describing language use, for example, and they are questions that the field of typology should also consider.

**Acknowledgements:** The authors gratefully acknowledge support from the UK Arts and Humanities Research Council (AH/N00924X/1).

## References

- Cameron, Richard. 1998. A variable syntax of speech, gesture, and sound effect: Direct quotations in Spanish. *Language Variation and Change* 10. 43–83. doi:10.1017/S0954394500001216.
- Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H. 2016. Depicting as a method of communication. *Psychological Review* 123(3). 324–347. doi:10.1037/rev0000026.



- Clark, Herbert H. & Richard Gerrig. 1990. Quotations as demonstrations. *Language* 66(4). 764–805. doi:10.2307/414729.
- Cormier, Kearsy, Sandra Smith & Zed Sevcikova. 2015. Rethinking constructed action. *Sign Language and Linguistics* 18(2). 167–204. doi:10.1075/sll.18.2.01cor.
- Cormier, Kearsy, Sandra Smith & Martine Zwets. 2013. Framing constructed action in BSL narratives. *Journal of Pragmatics* 55. 119–139. doi:10.1016/j.pragma.2013.06.002.
- Croft, William. 2016. Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology* 20(2). 377–393. doi:10.1515/lingty-2016-0012.
- D'Arcy, Alexandra. 2015. Quotation and advances in understanding syntactic systems. *Annual Review of Linguistics* 1(1). 43–61. doi:10.1146/annurev-linguist-030514-125220.
- Dingemanse, Mark. 2015. Ideophones and reduplication: Depiction, description, and the interpretation of repeated talk in discourse. *Studies in Language* 39. 946–970. doi:10.1075/sl.39.4.05din.
- Dingemanse, Mark. 2019. “Ideophone” as a comparative concept. In Kimi Akita & Prashant Pardeshi (eds.), *Ideophones, mimetics, and expressives*. Amsterdam: John Benjamins. 13–34.
- Djenar, Dwi Noverini, Michael Ewing & Howard Manns. 2018. *Style and intersubjectivity in youth interaction*. Berlin: Mouton de Gruyter.
- Enfield, Nicholas J. 2009. *The anatomy of meaning: Speech, gesture, and composite utterances*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511576737.
- Ferrara, Lindsay & Rolf P. Halvorsen. 2017. Depicting and describing meanings with iconic signs in Norwegian Sign Language. *Gesture* 16(3). 371–395. doi:10.1075/gest.00001.fer.
- Ferrara, Lindsay & Gabrielle Hodge. 2018. Language as description, indication, and depiction. *Frontiers in Psychology* 9. 716. doi:10.3389/fpsyg.2018.00716.
- Ferrara, Lindsay & Trevor Johnston. 2014. Elaborating who's what: A study of enactment and clause structure in Auslan (Australian Sign Language). *Australian Journal of Linguistics* 34(2). 193–215. doi:10.1080/07268602.2014.887405.
- Givón, Thomas. 2009. *The Genesis of syntactic complexity: Diachrony ontogeny neuro-cognition evolution*. Amsterdam/Philadelphia: John Benjamins.
- Goldin-Meadow, Susan & Diane Brentari. 2017. Gesture, sign and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences* 40. e46. doi:10.1017/S0140525X15001247.
- Haiman, John. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.
- Hodge, Gabrielle. 2014. Patterns from a signed language corpus: Clause-like units in Auslan (Australian sign language). Unpublished PhD thesis. Department of Linguistics, Macquarie University, Sydney. <http://hdl.handle.net/1959.14/323269>
- Hodge, Gabrielle & Trevor Johnston. 2014. Points, depictions, gestures and enactment: Partly lexical and non-Lexical signs as core elements of single clause-like units in Auslan (Australian Sign Language). *Australian Journal of Linguistics* 34. 262–291. doi:10.1080/07268602.2014.887408.
- Johnston, Trevor. 1996. Function and medium in the forms of linguistic expression found in a sign language. In William H. Edmondson & Ronnie B. Wilbur (eds.), *International review of sign linguistics*, vol. 1. 57–94. Mahwah, New Jersey: Lawrence Erlbaum.
- Johnston, Trevor & Adam Schembri. 2010. Variation, lexicalization and grammaticalization in signed languages. *Langage et société* 131(March). 19–35.



- Kendon, Adam. 2014. Semiotic diversity in utterance production and the concept of 'language'. *Philosophical Transactions of the Royal Society of London B Biological Sciences* 369. 20130293.
- Kockelman, Paul. 2005. The semiotic stance. *Semiotica* 157(1/4). 233–304. doi:10.1515/semi.2005.2005.157.1-4.233.
- Kusters, Annelies, Massimiliano Spotti, Ruth Swanwick & Elina Tapio. 2017. Beyond languages, beyond modalities: Transforming the study of semiotic repertoires. *International Journal of Multilingualism* 14. 219–232. doi:10.1080/14790718.2017.1321651.
- LaPolla, Randy J. 2003. Why languages differ: Variation in the conventionalization of constraints on inference. In David Bradley, Randy LaPolla, Boyd Michailovsky & Graham Thurgood (eds.), *Language variation: Papers on variation and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, 113–144. Canberra: Australian National University.
- Mazeland, Harrie. 2006. VAN as a quotative in Dutch: Marking quotations as a typification. In Tom Koole, Jacomine Nortier & Bert Tahitu (eds.), *Artikelen voor de 5e Sociolinguïstische Conferentie*, 354–365. Delft: Oburon.
- Metzger, Melanie. 1995. Constructed dialogue and constructed action in American Sign Language. In Ceil Lucas (ed.), *Sociolinguistics in deaf communities*, 255–271. Washington, DC: Gallaudet University Press.
- Okrent, Arika. 2002. A modality-free notion of gesture and how it can help us with the morpheme vs. gesture questions in sign language linguistics (or at least give us some criteria to work with). In Richard P. Meier, Kearsy Cormier & David Quinto-Pozos (eds.), *Modality and structure in signed and spoken languages*, 175–198. New York: Cambridge University Press.
- Park, Yujong. 2009. Interaction between grammar and multimodal resources: Quoting different characters in Korean multiparty conversation. *Discourse Studies* 11(1). 79–104. doi:10.1177/1461445608098499.
- Quinto-Pozos, David. 2007. Can constructed action be considered obligatory. *Lingua* 117. 1285–1314. doi:10.1016/j.lingua.2005.12.003.
- Sams, Jessie. 2010. Quoting the unspoken: An analysis of quotations in spoken discourse. *Journal of Pragmatics* 42. 3147–3160. doi:10.1016/j.pragma.2010.04.024.
- Schembri, Adam, Jordan Fenlon, Kearsy Cormier & Trevor Johnston. 2018. Sociolinguistic typology and sign languages. *Frontiers in Psychology* 9. 200. doi:10.3389/fpsyg.2018.00200.
- Sidnell, Jack. 2006. Coordinating gesture, talk, and gaze in reenactments. *Research on Language & Social Interaction* 39. 377–409. doi:10.1207/s15327973rlsi3904\_2.
- Spronck, Stef. 2017. Defenestration: Deconstructing the frame-in relation in Ungarinyin. *Journal of Pragmatics* 114. 104–133. doi:10.1016/j.pragma.2017.03.016.
- Stec, Kashmiri, Mike Huiskes & Gisela Redeker. 2016. Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics* 104. 1–17. doi:10.1016/j.pragma.2016.07.008.
- Tannen, Deborah. 1989. *Talking voices*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Vološinov, Valentin. N. 1973. *Marxism and the philosophy of language*. New York/London: Seminar Press.