Target article: The coming of age of sign language and gesture studies
Authors: Susan Goldin-Meadow & Diane Brentari
Word counts:

- Abstract: 56 words
- Main paper: 999 words
- Figure caption: 101
- References: 112 words
- Total word count: 1268 words

Where does (sign) language begin?

Iris Berent[1] *and Amanda Dupuis*
Northeastern University

Department of Psychology
Northeastern University
125 Nightingale Hall
Boston, MA 02215

[1] Corresponding author: i.berent@neu.edu, 617 373 4033

Abstract. Goldin-Meadow and Brentari outline several criteria for delineating the boundaries between (discrete) signs and (continuous) gestures. However, the complex links between linguistics forms and their phonetic realizations defy such heuristics. A systematic exploration of language structure by mouth and by hand may help get us closer to answering the important challenge outlined in this target paper.

Where does (sign) language begin and where do (nonlinguistic) gestures end is a critical theoretical question that is central to an account of sign language, specifically, and the language faculty, broadly. Goldin-Meadow and Brentari (GM&B)'s target article makes important strides towards its resolution. At the theoretical level, the authors convincingly demonstrate that linguistic forms and gestures exhibit stark differences that are suggestive of distinct computational origins, and these distinctions are evident irrespective of language modality—manual or aural. This conclusion is significant, as it shows that the differences between manual and spoken language might be smaller than what meets the eye/ear. Methodologically, GM&B also outline several criteria for demarcating the boundaries between sign and gesture.

We applaud the authors' theoretical efforts and pioneering empirical work. However, it is important to recognize that their criteria for distinguishing signs and gestures are merely useful empirical heuristics--they will not suffice in and of themselves to define the boundaries of the language faculty.

GM&B seek to distinguish signs and gestures by contrasting their phonetic forms, meanings and pragmatic functions. Signs, in their view, exhibit discrete phonetic form whereas gestures are continuous; signs' meanings are at least partly conventional and arbitrary whereas gestures convey imagistic information using non-arbitrary means, hence, they are largely independent of experience with sign language. Finally signs and gestures differ pragmatically inasmuch as they can convey different (and even contradictory) aspects of thought (e.g., during problem solving).

Although these three criteria can help identify (nonlinguistic) gestures, their utility for defining linguistic forms is less clear. Critically, these difficulties are expected even if signs and gestures do in fact originate from distinct computational mechanisms—an algebraic grammar vs. an analog conceptual interface, respectively.

Considering first the phonetic criteria, the links between discrete linguistic categories and their phonetic realizations are far from transparent. While analog nonlinguistic computations (e.g., for gestures) are likely to give rise to "phonetic" gradiance, gradiance could also result from the realization of grammatical categories that are discrete and abstract. To use an example from spoken language, *scenery* and *chicanery* are each equally good members of the *Noun* category, as these exemplars are equally admissible to grammatical computations that apply to the category as a whole (e.g., regular inflection). But at the phonetic level, these exemplars will likely acquire gradient phonetic manifestations—high frequency forms, for instance, are more likely to undergo schwa reduction (e.g., *scenery*→*scenry*) than low frequency forms (e.g., *chicanery*→*chicanry*; Bybee, 2002). Accordingly, a phonetic inspection of these exemplars may not necessarily inform us of their grammatical status.

Extending this logic to GM&B's own example from signs, the fact that the phonetic realization of verb agreement (i.e., height in signing space) varies continuously depending on the addressee (adult or child) does not negate the possibility that the categories that inform syntactic computations are discrete and abstract, free of that analog information. Similarly, the gradient phonetic implementation of movement and location does not necessarily inform phonological processes, so phonetic gradiance is entirely consistent with the possibility that the phonological grammar of sign languages are algebraic and abstract (Berent, Dupuis, & Brentari, 2014). The disyllabic noun *seat*, for instance, is likely represented algebraically, as fully reduplicative (i.e., as XX), even if the location and movement features of its two syllables are phonetically distinct, and these differences are noticeable by signers in some other context (e.g., phonetic categorization). Accordingly, the phonetic realization of a manual form cannot transparently indicate its mental representation by the grammar. While gestures are likely to take continuous phonetic forms, phonetic gradiance might also realize linguistic signs that are discrete and abstract.

In fact, judging by the literature from spoken language, any given sensory form may well acquire multiple representations at different levels of analysis—the dual percepts of speech analogs (as either linguistic speech, or nonlinguistic nonspeech) attests to this fact (Remez, Pardo, Piorkowski, & Rubin, 2001). Furthermore, speakers of different languages (e.g., Russian vs. English) demonstrably project their linguistic knowledge to the perception of nonlinguistic stimuli (i.e., nonspeech)—the better formed the stimulus in their native language, the more likely its perception as speech (Berent, Balaban, Lennertz, & Vaknin-Nusbaum, 2010). These observations are significant because they suggest that the functional role of a given input—as linguistic sign (spoken or manual) or nonlinguistic element (e.g.,

gesture/nonspeech) is determined (at least in part) by grammatical constraints, and consequently, it is unpredictable solely from its phonetic form.

Experience-dependence (e.g., differences between signers and nonsigners) may likewise fail to reveal the status of a stimulus as "linguistic". GM&B show that the silent gesturing of nonsigners has many characteristics of grammatical signs. Other evidence suggests that nonsigners' representation of signs relies not only on visual strategies but also on shared grammatical constraints. For example, our past research (Berent, Dupuis, & Brentari, 2013) shows that signers and nonsigners both define signed syllables by the number of sonority peaks (i.e., movement)—a principle that is likely universal and amodal. Critically, these biases are linguistic, rather than merely visual, as nonsigners selectively apply them to syllables, but not to morphemes (see Figure 1). Furthermore, while nonsigners readily learn this UG-consistent regularity (syllables are defined by movements; morphemes by handshape), they are unable to learn the reverse (syllables are defined by handshapes; morphemes by movements). The potential for some linguistic biases to extend across modality and linguistic experience suggests caution in applying these criteria in the definition of signs.

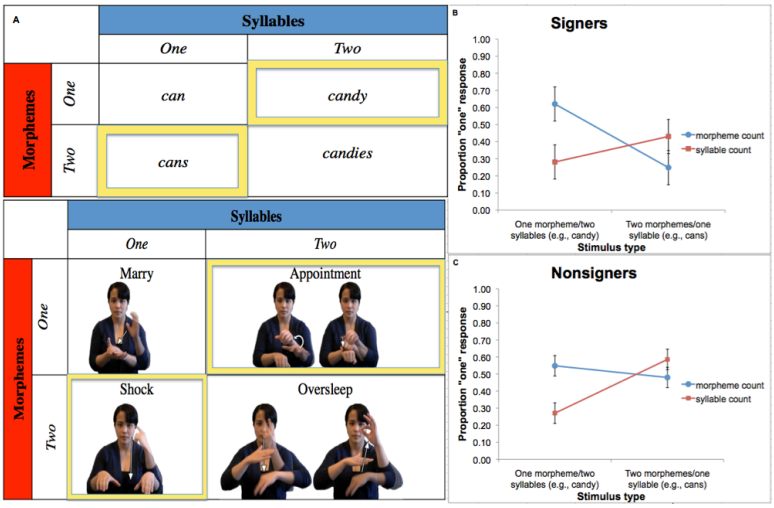

Figure 1. Amodal restrictions on syllable structure. (a) Signed and spoken languages contrast syllables and morphemes; syllables are defined by phonological elements that correspond to energy peaks—either vowels (in English) or movements (in American Sign Language). Furthermore, signers (b) and nonsigners (c) can track the number of syllables and novel signs. Accordingly, when presented incongruency between the number of syllables and morphemes (see the highlighted cells), people shift their response (shown as the proportion of "one" responses, either one syllable or one morpheme) depending on whether they are asked to count syllables or morphemes. Data from Berent et al., (2013).

Where, then, does (sign) language begin? We do not have a hard and fast solution to this question. However, it is important recognize that the identification of linguistic inputs as such might be partly the product of linguistic computations rather than sensory and motor mechanisms alone. We thus believe it might be useful to complement GM&B's empirical heuristics by a deductive approach that departs from a formal account of the language faculty and experimentally compares its implementation across modalities. A systematic exploration of language structure by mouth and by hand may help get us closer to answering the important challenge outlined by this target paper.

References

Berent, I., Balaban, E., Lennertz, T., & Vaknin-Nusbaum, V. (2010). Phonological universals constrain the processing of nonspeech. *Journal of Experimental Psychology: General, 139*, 418-435.

Berent, I., Dupuis, A., & Brentari, D. (2013). Amodal aspects of linguistic design. *PLoS ONE, 8*.

Berent, I., Dupuis, A., & Brentari, D. (2014). Phonological reduplication in sign language: Rules rule. *Frontiers in Language Sciences, 5*, 560.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change, 14*, 261-290.

Remez, R. E., Pardo, J. S., Piorkowski, R. L., & Rubin, P. E. (2001). On the bistability of sine wave analogues of speech. *Psychological science, 12*, 24-29.