Published in final edited form as:

Behav Brain Sci. 2016 January ; 39: e70. doi:10.1017/S0140525X1500076X.

Consequences of the Now-or-Never bottleneck for signed versus spoken languages

Karen Emmorey

School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, CA 92120, kemmorey@mail.sdsu.edu, <http://emmoreylab.sdsu.edu/>

Abstract

Signed and spoken languages emerge, change, are acquired, and are processed under distinct perceptual, motor, and memory constraints. Therefore, the Now-or-Never bottleneck has different ramifications for these languages, which are highlighted in this commentary. The extent to which typological differences in linguistic structure can be traced to processing differences provides unique evidence for the claim that structure is processing.

Christiansen & Chater (C&C) make it clear that the consequences of the Now-or-Never bottleneck for language are not speech-specific. This commentary highlights how and why signed and spoken languages respond differently to the limitations imposed by the bottleneck. C&C argue that the Now-or-Never bottleneck arises from general principles of perceptuo-motor processing and memory, and both have different properties for visual-manual and aural-oral languages, which lead to adaptations and preferences that are specific to each language type.

The vocal articulators (lips, tongue, larynx) are smaller and quicker than the hands and arms, and the auditory system is generally more adept at temporal processing than the visual system, which is better at spatial processing. These perceptual and motoric differences exert distinct pressures and affordances when solving the problems presented by the Now-or-Never bottleneck. As a consequence, signed and spoken languages prefer different chunking strategies for structuring linguistic information. At the phonological level, spoken languages prefer what could be called *serial* chunking, whereas signed languages prefer *spatial* chunking. For example, for spoken languages, single-segment words are rare and multisegment words are common, but the reverse pattern holds for sign languages (Brentari 1998). Oversimplifying here, consonants and vowels constitute segment types for speech, while locations and movements constitute segment types for sign (e.g., Sandler 1986). Single-segment spoken words are rare because they are extremely short and generally limited to the number of vowels in the language. Single consonants violate the Possible-Word Constraint (Norris et al. 1997), which also applies to sign language (Orfanidou et al. 2010). Single-segment signs are not problematic because other phonological information – for example, hand configuration – can be produced (and perceived) simultaneously with a large number of possible single location or movement segments. Multisegment (> three) and multisyllabic signs are rare in part because the hands are relatively large and slow articulators, and this limits the number of serial segments that can be quickly chunked and passed on to the lexical level of representation.

Distinct preferences for serial versus spatial chunking are also found at the morphological level. Spoken languages show a general preference for linear affixation (specifically, suffixation) over nonconcatenative processes such as reduplication or templatic morphology (Cutler 1985). In contrast, linear affixation (particularly for inflectional morphology) is rare across sign languages, and simultaneous, nonconcatenative morphology is the norm. Aronoff et al. (2005) attributed the paucity of linear morphology to the youth of sign languages but acknowledged that processing constraints imposed by modality also shape this preference (Emmorey 1995). Specifically, the ability of the visual system to process spatially distributed information in parallel, the slow articulation rate of the hands, and limits on working memory all conspire to induce sign languages to favor simultaneous over sequential morphological processes. In fact, when the linear morphology of a spoken language is implemented in the visual-manual modality, as in Manually Coded English (MCE), deaf children who have no exposure to a natural sign language spontaneously create simultaneous morphology to mark verb arguments (Supalla 1991). In addition, linear manual suffixes in MCE are often incorrectly analyzed by children as separate signs because prosodically and perceptually they do not pattern like bound morphemes (Supalla & McKee 2002).

Although the architecture of the memory system is parallel for signed and spoken languages (Wilson & Emmorey 1997; 1998), immediate memory for sequences of items has consistently been found to be superior for speech (Bellugi et al. 1975, *inter alia*). Hall and Bavelier (2009) demonstrated that the serial span discrepancy between speech and sign arises during perception and encoding, but not during recall, where sign actually shows an advantage (possibly because visual feedback during signing does not interfere with the memory store, unlike auditory feedback during speaking; Emmorey et al. 2009). The source of these differences is still unclear, but the short-term memory capacity for sign (4–5 items) is typical of a variety of types of memory (Cowan 2000), and thus what needs to be explained is why the memory capacity for speech is unusually high.

Because sign languages emerge, change, are acquired, and are processed under distinct memory and perceptuo-motor constraints, they provide an important testing ground for C&C's controversial proposals that learning to process is learning the grammar and that linguistic structure is processing history. Typological differences between the structure of signed and spoken languages may be particularly revealing. Can such structural differences be explained by distinct processing adaptations to the Now-or-Never bottleneck? For example, given the bottleneck pressures, one might expect duality of patterning to emerge quickly in a signed language, but recent evidence suggests that it may not (Sandler et al. 2011). Could this be because the visual-manual and auditory-oral systems are “lossy” in different ways or because chunking processes differ between modalities? Given C&C's claim that “there is no representation of grammatical structure separate from processing” (sect. 6.2, para. 6), it is critical to determine whether the differences – and the commonalities – between signed and spoken languages can be traced to features of processing.