(virtually) instantaneous (Chomsky 1975; 1980; 1986; 2012), C&C propose that "the Now-or-Never bottleneck requires that language acquisition be viewed as a type of skill learning, such as learning to drive, juggle, play the violin, or play chess. Such skills appear to be learned through *practicing* the skill, using online feedback during the practice itself …" (sect. 4, para. 4). This view integrates language naturally within cognition and does not require the postulation of domain-specific cognitive modules. Additionally, C&C's account casts doubt on Chomsky's claim that the fact that we frequently talk silently to ourselves supports his view that the function of language is not communication (e.g., Chomsky 2000; 2002; 2012). A more parsimonious explanation would assume that frequent internal monologues arose from the habitual "practice" (fine-tuning by [silently] doing) of language learning. C&C argue that "we should expect the exploitation of memory to require 'replaying' learned material, so that it can be reprocessed" (sect. 4.1, para. 5). They cite substantial neuroscientific evidence that such replay occurs and propose that dreaming may have a related function. Given that especially the integration of available information across different types and levels of abstraction and the anticipation of responses might require more practice than the motor-execution of (audible) speech, silent self-conversation might initially provide an additional medium for language learning. Later in life, such internal monologue could be recruited to the function Chomsky envisioned. Future research could uncover at what age children begin using internal monologue, to what degree second-language acquisition is assisted by learners switching their internal monologue from L1 to L2, and whether the lack of internal monologue (e.g., Grandin 2005) has negative effects on fluency in production.

Although C&C's account offers an attractive language acquisition model, it seems to create a paradox for language evolution. C&C argue that there are strong pressures toward simplification and reduction. For example, when a very simple artificial toy language was simulated, it "collapsed into just a few different forms that allowed for systematic, albeit semantically underspecified, generalization … In natural language, however, the pressure toward reduction is normally kept in balance by the need to maintain effective communication" (sect. 5, para. 4). This observation raises the following problem: For an existing, fairly complex system, simplification may indeed lead to the kinds of changes C&C discuss (e.g., that "chunks at each level of linguistic structure – discourse, syntax, morphology, and phonology – are potentially subject to reduction" [sect. 5, para. 5]). But in this view there is a strong pressure toward simplification and virtually no possibility of increasing complexity. Yet it is not clear why the language of our distant ancestors would have been more complex than or at least as complex as modern languages. It has been argued convincingly that the complexity of grammar actually needed to support most daily activities of humans living in complex contemporary societies is substantially less than that exhibited by any contemporary human language (Gil 2009, p. 19), and it seems implausible that existing language complexity is functionally motivated.

If the Now-or-Never bottleneck has the power C&C attribute to it, it must have constrained language learning and use for our distant ancestors in the same way as it does for us. Presumably these ancestors had cognitive capacities that were not superior to ours, and their culture would have imposed even fewer demands for linguistic complexity than contemporary culture. So how could they have evolved a highly complex language system that in turn could be reduced to provide the cognitive foundation for grammaticalization? C&C suggest analogies between language and other cognitive processes (e.g., vision). This is problematic because the visual system evolved to perceive objects that exist independently of this system. On purely naturalist accounts, languages have no existence independent of human brains or human culture. Therefore, both the cognitive abilities underwriting linguistic competence and the language that is learned must have evolved. Decades ago it was suggested that many of the problems that bedevil Chomskyan linguistics can be eliminated if one adopts linguistic Platonism and draws a distinction between the knowledge speakers have of their language and the languages that speakers have knowledge of. Platonism considers as distinct (1) the study of semantic properties and relations like ambiguity, synonymy, meaningfulness, and analyticity, and (2) the study of the neuropsychological brain-states of a person who has acquired knowledge about these semantic properties (e.g., Behme 2014a; Katz 1984; 1996; 1998; Katz & Postal 1991; Neef 2014; Postal 2003; 2009). In such a view, languages and brains that have acquired knowledge of languages are two distinct ontological systems.

In addition to eliminating many problems for contemporary linguistics, such a view also might resolve the language evolution paradox because languages have an independent existence, and only human cognitive capacity evolves. It might be argued that the epistemology of linguistic Platonism is hopeless. Although this is not the place to defend linguistic Platonism, one should remember that in mathematics it is widely accepted that the number systems exist independently of human brains and human culture, and are discovered, just as are other objects of scientific discovery. It has been argued that if one accepts the possibility of mathematical realism, there is no a priori reason to reject the possibility of linguistic realism (e.g., Behme 2014b; Katz 1998). Before rejecting linguistic Platonism out of hand, one ought to remember that

> For psychology, AI, and the related cognitive sciences, the question of what a grammar is a theory of is important because its answer can resolve troublesome issues about where the linguist's work ends and the cognitive scientist's begins. A Platonist answer to this question would clearly divide linguistics and cognitive sciences so that the wasteful and unnecessary quarrels of the past can be put behind us. (Katz 1984, p. 28)

ACKNOWLEDGMENTS

I thank Vyvyan Evans, David Johnson, and Paul Postal for comments on earlier drafts and accept full responsibility for remaining errors.

Socio-demographic influences on language structure and change: Not all learners are the same

doi:10.1017/S0140525X15000710, e66

Till Bergmann,[a] Rick Dale,[a] and Gary Lupyan[b]

[a]*Cognitive and Information Sciences, University of California, Merced, Merced, CA 95343;* [b]*Department of Psychology, University of Wisconsin–Madison, Madison, WI 53706.*

tbergmann@ucmerced.edu www.tillbergmann.com
rdale@ucmerced.edu http://cognaction.org/rick/
glupyan@gmail.com http://sapir.psych.wisc.edu/

Abstract: The Now-or-Never bottleneck has important consequence for understanding why languages have the structures they do. However, not addressed by C&C is that the bottleneck may interact with *who* is doing the learning: While some languages are mostly learned by infants, others have a large share of adult learners. We argue that such socio-demographic differences extend and qualify C&C's thesis.

We wholeheartedly agree with Christiansen & Chater (C&C) that "acquiring a language is learning to process" (sect. 5, para. 3) and that "there is no representation of grammatical structure separate from processing" (sect. 6.2, para. 6). We also agree with C&C's more general thesis that the structure of language cannot be understood without taking into account the constraints and biases of the language learners and users. Although the Now-or-Never

cognitive bottleneck is an unavoidable constraint on language comprehension and production, fully understanding its consequences requires taking into account socio-demographic realities, namely *who* is doing the language learning.

C&C write that "Language will be shaped by the linguistic patterns learners find easiest to acquire and process" (sect. 5, para. 3), but what is *easiest* may importantly depend on *who* is doing the learning. Some languages are learned exclusively by infants and used in small, culturally homogeneous communities. For example, half of all languages have fewer than 7,000 speakers. Other languages have substantial populations of non-native speakers and are used in large, culturally and linguistically heterogeneous communities. For example, at present, about 70% of English speakers are non-native speakers (Gordon 2005).

The socio-demographic niche in which a language is learned and used can influence its grammar insofar as different kinds of learners have different learning biases. Languages with many adult learners may adapt to their socio-demographic niche by eschewing features difficult for adults to process. Indeed, as Lupyan and Dale (2010) have shown in an analysis of more than 2,000 languages, languages spoken in larger and more diverse communities (those that tend to have more non-native learners) have simpler morphologies and fewer obligatory markings (see also Bentz & Winter 2013). In contrast, languages used in a socio-demographic niche containing predominantly infant learners tend to have many more obligatory markings – for example, they are more likely to encode tense, aspect, evidentiality, and modality as part of the grammar, and to have more complex forms of agreement (Dale & Lupyan 2012; see also Dahl 2004; McWhorter 2001; Trudgill 2011).

Such influences of the socio-demographic environment on language structure are important caveats to C&C's thesis because the Now-or-Never bottleneck, although present in all learners, depends on the knowledge that a learner brings to the language-learning task.

On C&C's account, successful language processing depends on recoding the input into progressively higher-level (more abstract) chunks. As an analogy, C&C give the example of how remembering strings of numbers is aided by chunking (re-representing) numbers as running times or dates (sect. 2, para. 7). But, of course, this recoding is only possible if the learner knows about reasonable running times and the format of dates. The ability to remember the numbers depends on the ability to chunk them, and the ability to chunk them depends on prior knowledge.

In the case of language learning, recoding of linguistic input is "achieved by applying the learner's *current* model of the language" (sect. 4.1, para. 3) and further constrained by memory and other domain-general processes. But both the learner's language model and domain-general constraints vary depending on *who* the learner is.

Infants come to the language-learning task with a less developed memory and ability to use pragmatic and other extralinguistic cues to figure out the meaning of an utterance. As a consequence, the Now-or-Never bottleneck is strongly in place. The language adapts through increased grammaticalization that binds units of meaning more tightly, thereby increasing redundancy. For example, grammatical gender of the sort found in many Indo-European languages increases redundancy by enforcing agreement of nouns, adjectives, and pronouns, making one more predictable from the other and – arguably – reducing the memory load required for processing the utterances.

Adults come to the language-learning task with more developed memories, and ability for pragmatic inference, but at the same time they are biased by pre-existing chunks that may interfere with chunks that most efficiently convey the meaning in the new language. The greater memory capacities and ability to use contextual and other pragmatic cues to infer meanings, may relax the Now-or-Never bottleneck, nudging grammars toward morphological simplification with its accompanying decrease in obligatory markings (i.e., decrease in redundancy) and increase in compositionality (Lupyan & Dale 2015).

This reasoning helps explain *how* the Now-or-Never bottleneck can create "obligatorification" (sect. 5, para. 8) and also why some languages have more obligatory markings than other languages.

In summary, although we agree with C&C that "multiple forces influence language change in parallel" (sect. 5, para. 9), we emphasize the force constituted by the learning community. Languages adapt to the specific (cognitive) learning constraints and communicative needs of the learners and speakers.

Now or … later: Perceptual data are *not* immediately forgotten during language processing

doi:10.1017/S0140525X15000734, e67

Klinton Bicknell,[a] T. Florian Jaeger,[b,c,d] and Michael K. Tanenhaus[b,d]

[a]*Department of Linguistics, Northwestern University, Evanston, IL 60208-0854;* [b]*Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627-0268;* [c]*Department of Computer Science, University of Rochester, Rochester, NY 14627-0226;* [d]*Department of Linguistics, University of Rochester, Rochester, NY 14627-0096.*

kbicknell@northwestern.edu fjaeger@mail.bcs.rochester.edu
mtan@mail.bcs.rochester.edu
http://faculty.wcas.northwestern.edu/kbicknell/
http://www.bcs.rochester.edu/people/fjaeger/
https://www.bcs.rochester.edu/people/mtan/mtan.html

Abstract: Christiansen & Chater (C&C) propose that language comprehenders must immediately compress perceptual data by "chunking" them into higher-level categories. Effective language understanding, however, requires maintaining perceptual information long enough to integrate it with downstream cues. Indeed, recent results suggest comprehenders do this. Although cognitive systems are undoubtedly limited, frameworks that do not take into account the tasks that these systems evolved to solve risk missing important insights.

Christiansen & Chater (C&C) propose that memory limitations force language comprehenders to compress perceptual data immediately, forgetting lower-level information and maintaining only higher-level categories ("chunks"). Recent data from speech perception and sentence processing, however, demonstrate that comprehenders can maintain fine-grained lower-level perception information for substantial durations. These results directly contradict the central idea behind the Now-or-Never bottleneck. To the extent that the framework allows them, it risks becoming so flexible that it fails to make substantive claims. On the other hand, these results are predicted by existing frameworks, such as bounded rationality, which are thus more productive frameworks for future research. We illustrate this argument with recent developments in our understanding of a classic result in speech perception: categorical perception.

Initial results in speech perception suggested that listeners are insensitive to fine-grained within-category differences in voice onset time (VOT, the most important cue distinguishing voiced and voiceless stop consonants, e.g., "b" versus "p" in *bill* versus *pill*), encoding only whether a sound is "voiced" or "voiceless" (Liberman et al. 1957). Subsequent work demonstrated sensitivity to within-category differences (Carney et al. 1977; Pisoni & Tash 1974), with some findings interpreted as evidence this sensitivity rapidly decays (e.g., Pisoni & Lazarus 1974). Such a picture is very similar to the idea behind Chunk-and-Pass: Listeners rapidly chunk phonetic detail into a phoneme, forgetting the subcategorical information in the process.

Although it may perhaps be intuitive, given early evidence that perceptual memory is limited (Sperling 1960), such discarding of subcategorical information would be surprising from the perspective