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.

Linguistics, cognitive psychology, and the Now-or-Never bottleneck

Ansgar D. Endress[a] and Roni Katzir[b]

[a]Department of Psychology, City University London, London EC1V 0HB, United Kingdom; [b]Department of Linguistics and Sagol School of Neuroscience, Tel Aviv University, Ramat Aviv 69978, Israel.
ansgar.endress.1@city.ac.uk   rkatzir@post.tau.ac.il

Abstract: Christiansen & Chater (C&C)'s key premise is that "if linguistic information is not processed rapidly, that information is lost for good" (sect. 1, para. 1). From this "Now-or-Never bottleneck" (NNB), C&C derive "wide-reaching and fundamental implications for language processing, acquisition and change as well as for the structure of language itself" (sect. 2, para. 10). We question both the premise and the consequentiality of its purported implications.

Problematic premises. Christiansen & Chater (C&C) base the Now-or-Never bottleneck (NNB) on the observation that sensory memory disappears quickly in explicit memory tasks. We note, first, that not all forms of explicit memory are short-lived. For example, children remember words encountered once after a month (Carey & Bartlett 1978; Markson & Bloom 1997). More important, it is by no means clear that explicit memory is the (only) relevant form of memory for language processing and acquisition, nor how quickly other forms of memory decay. For example, the perceptual learning literature suggests that learning can occur even in the absence of awareness of the stimuli (Seitz & Watanabe 2003; Watanabe et al. 2001) and sometimes has long-lasting effects (Schwab et al. 1985). Similarly, visual memories that start decreasing over a few seconds can be stabilized by presenting items another time (Endress & Potter 2014). At a minimum, then, such memory traces are long-lasting enough for repeated exposure to have cumulative learning effects.

Information that is not even perceived is thus used for learning and processing, and some forms of memory do not disappear immediately. Hence, it is still an open empirical question whether poor performance in explicit recall tasks provides severe constraints on processing and learning.

We note, in passing, that even if relevant forms of memory were short-lived, this would not necessarily be a bottleneck. Mechanisms to make representations last longer – such as self-sustained activity – are well documented in many brain regions (Major & Tank 2004), and one might assume that memories can be longer-lived when this is adaptive. Short-lived memories might thus be an adaptation rather than a bottleneck (e.g., serving to reduce information load for various computations).

Problematic "implications." C&C use the NNB to advance the following view: Language is a skill (specifically, the skill of parsing predictively); this skill is what children acquire (rather than some theory-like knowledge); and there are few if any restrictions on linguistic diversity. C&C's conclusions do not follow from the NNB and are highly problematic. Below, we discuss some of the problematic inferences regarding processing, learning, and evolution.

Regarding processing, C&C claim that the NNB implies that knowledge of language is the skill of parsing predictively. There is indeed ample evidence for a central role for prediction in parsing (e.g., Levy 2008), but this is not a consequence of the NNB: The advantages of predictive processing are orthogonal to the NNB, and, even assuming the NNB, processing might still occur element by element without predictions. C&C also claim that the NNB implies a processor with no explicit representation of syntax (other than what can be read off the parsing process as a trace). It is unclear what they actually mean with this claim, though. First, if C&C mean that the parser does not construct full syntactic trees but rather produces a minimum that allows semantics and phonology to operate, they just echo a view discussed by Pulman (1986) and others. Although this view is an open possibility, we do not see how it follows from the NNB. Second, if C&C mean that the NNB implies that parsing does not use explicit syntactic knowledge, this view is incorrect: Many parsing algorithms (e.g., LR, Earley's algorithm, incremental CKY) respect the NNB by being incremental and not needing to refer back to raw data (they can all refer to the result of earlier processing instead) and yet make reference to explicit syntax. Finally, we note that prediction-based, parser-only models in the literature that do not incorporate explicit representations of syntactic structure (e.g., Elman 1990; McCauley & Christiansen 2011) fail to explain why we can recognize unpredictable sentences as grammatical (e.g., *Evil unicorns devour xylophones*).

Regarding learning, C&C claim that the NNB is incompatible with approaches to learning that involve elaborate linguistic knowledge. This, however, is incorrect: The only implication of the NNB for learning is that if memory is indeed fleeting, any learning mechanism must be online rather than batch, relying only on current information. But online learning does not rule out theory-based models of language in any way (e.g., Börschinger & Johnson 2011). In fact, some have argued that online variants of theory-based models provide particularly good approximations to empirically observed patterns of learning (e.g., Frank et al. 2010).

Regarding the evolution of language (which they conflate with the biological evolution of language), C&C claim that it is item-based and gradual, and that linguistic diversity is the norm, with few if any true universals. However, how these claims might follow from the NNB is unclear, and C&C are inconsistent with the relevant literature. For example, language change has been argued to be abrupt and nonlinear (see Niyogi & Berwick 2009), often involving what look like changes in abstract principles rather than concrete lexical items. As for linguistic diversity, C&C repeat claims from Christiansen and Chater (2008) and Evans and Levinson (2009), but those works ignore the strongest typological patterns revealed by generative linguistics. For example, no known language allows for a single conjunct to be displaced in a question (Ross 1967): We might know that Kim ate *peas and*

something yesterday and wonder what that *something* is, but in no language can we use a question of the form *What did Kim eat peas and yesterday?* to inquire about it. Likewise, in *Why did John wonder who Bill hit?*, one can only ask about the cause of the wondering, not of the hitting (see Huang 1982; Rizzi 1990). Typological data thus reveal significant restrictions on linguistic diversity.

Conclusion. Language is complex. Our efforts to comprehend it are served better by detailed analysis of the cognitive mechanisms at our disposal than by grand theoretical proposals that ignore the relevant psychological, linguistic, and computational distinctions.

Is Now-or-Never language processing good enough?

Fernanda Ferreira[a] and Kiel Christianson[b]

[a]Department of Psychology, University of California, Davis, Davis, CA 95616;
[b]College of Education, University of Illinois, Urbana–Champaign, Champaign, IL 61820.
fferreira@ucdavis.edu
http://psychology.ucdavis.edu/people/fferreir
kiel@illinois.edu
http://epl.beckman.illinois.edu/

Abstract: Christiansen & Chater's (C&C's) Now-or-Never bottleneck framework is similar to the Good-Enough Language Processing model (Ferreira et al. 2002), particularly in its emphasis on sparse representations. We discuss areas of overlap and review experimental findings that reinforce some of C&C's arguments, including evidence for underspecification and for parsing in "chunks." In contrast to Good-Enough, however, Now-or-Never does not appear to capture misinterpretations or task effects, both of which are important aspects of comprehension performance.

Christiansen & Chater (C&C) offer an intriguing proposal concerning the nature of language, intended to explain fundamental aspects of language comprehension, production, learning, and evolution. We agree with the basic framework, and indeed we have offered our own theoretical approach, Good-Enough (GE) Language Processing, to capture many of the phenomena discussed in the target article, particularly those relating to both online and offline comprehension. In this commentary, we hope to expand the discussion by pointing to some of these connections and highlighting additional phenomena that C&C did not discuss but that reinforce some of their points. In addition, however, we believe the GE model is better able to explain important aspects of language comprehension that C&C consider, as well as several they leave out. Of course, no single article could be comprehensive when it comes to a field as broad and active as this one, but we believe a complete theory of language must ultimately have something to say about these important phenomena, and particularly the content of people's interpretations.

We begin, then, with a brief review of the GE approach (Ferreira et al. 2002). The fundamental assumption is that interpretations are often shallow and sometimes inaccurate. This idea that interpretations are shallow and underspecified is similar to C&C's suggestion that the comprehension system creates chunks that might not be combined into a single, global representation. In their model, this tendency arises from memory constraints that lead the system to build chunks at increasingly abstract levels of representation. As evidence for this assumption regarding underspecified representations, C&C might have

discussed our work demonstrating that ambiguous relative clauses are often not definitively attached into the matrix structure if a failure to attach has no interpretive consequences (Swets et al. 2008; cf. Payne et al. 2014). Very much in line with C&C, Swets et al. observed that people who are asked detailed comprehension questions probing their interpretation of the ambiguous relative clause make definitive attachments, but those asked only shallow questions about superficial features of the sentence seem to leave the relative clause unattached – that is, they underspecify. This finding fits neatly with C&C's discussion of "right context effects," where here "right context" can be broadly construed to mean the follow-on comprehension question that influences the interpretation constructed online. An important difference, however, emerges as well, and here we believe the GE framework has some advantages over Now-or-Never as a broad model of comprehension: Our framework predicts that the language user's task will have a strong effect on the composition of "chunks" and the interpretation created from them (cf. Christianson & Luke 2011; Lim & Christianson 2015). We have reported these results in production as well, demonstrating that the extent to which speaking is incremental depends on the processing demands of the speaking task (Ferreira & Swets 2002). Given the importance of task effects in a range of cognitive domains, any complete model of language processing must include mechanisms for explaining how they arise.

Moreover, the idea that language processing proceeds chunk-by-chunk is not novel. C&C consider some antecedents of their proposal, but several are overlooked. For example, they argue that memory places major constraints on language processing, essentially obligating the system to chunk and interpret as rapidly as possible (what they term "eager processing"). This was a key motivation for Lyn Frazier's original garden-path model (Frazier & Rayner 1982) and the parsing strategies known as minimal attachment and late closure: The parser's goal is to build an interpretation quickly and pursue the one that emerges first rather than waiting for and considering multiple alternatives. This, too, is part of C&C's proposal – that the parser cannot construct multiple representations at the same level in parallel – but the connections to the early garden-path model are not mentioned, and the incompatibility of this idea with parallel models of parsing is also not given adequate attention. Another example is work by Tyler and Warren (1987), who showed that listeners form unlinked local phrasal chunks during spoken language processing and who conclude that they could find no evidence for the formation of a global sentence representation. Thus, several of these ideas have been part of the literature for many years, and evidence for them can be found in research motivated from a broad range of theoretical perspectives.

Perhaps the most critical aspect of comprehension that C&C's approach does not capture is meaning and interpretation: C&C describe an architecture that can account for some aspects of processing, but their model seems silent on the matter of the content of people's interpretations. This is a serious shortcoming given the considerable evidence for systematic misinterpretation (e.g., Christianson et al. 2001; 2006; Patson et al. 2009; van Gompel et al. 2006). In our work, we demonstrated that people who read sentences such as *While Mary bathed the baby played in the crib* often derive the interpretation that Mary bathed the baby, and they also misinterpret simple passives such as *The dog was bitten by the man* (Ferreira 2003). These are not small tendencies; the effects are large, and they have been replicated in numerous studies across many different labs. For C&C, these omissions are a lost opportunity because these results are consistent with their proposed architecture. For example, misinterpretations of garden-path sentences arise in part because the parser processes sentences in thematic chunks and fails to reconcile the various meanings constructed online. Recently, we demonstrated that the misinterpretations are attributable to a failure to "clean up" the interpretive consequences of creating these chunks (Slattery et al. 2013), a finding compatible with C&C's