Chapter 5

Language intertwined across multiple timescales: Processing, acquisition and evolution

Morten H. Christiansen
Department of Psychology, Cornell University, Ithaca, NY, USA
Centre for Interacting Minds, Aarhus University, Denmark

Theories of language invoke different types of causal dependencies to explain a variety of linguistic phenomena, ranging from typological patterns (e.g., 'verb-final languages tend to have postpositions,' Greenberg 1966) to psycholinguistic regularities (e.g., 'hearing a passive construction increases the likelihood of producing one,' Bock 1986). Several chapters in this volume provide important insights into such dependencies across a variety of domains (see for example, chapters by Cristofaro, Culbertson, Dediu, Hyman, and Rice). This chapter, however, concerns itself with a different kind of dependency: the fundamental theoretical interdependencies between different timescales of language, from processing to acquisition to evolution.

In the mainstream generative grammar tradition, possible interdependencies between language processing, acquisition and evolution are rarely explored (though see Pinker 1994; Jackendoff 2002). This is likely a consequence of Chomsky's methodological dictums that the study of language proper should be separated from how it is used and processed (Chomsky 1965), acquired over development (Chomsky 1975), and how it evolved (Chomsky 2005). Christiansen & Chater (2016a) refer to the theoretical impact of these methodological dictums as "Chomsky's hidden legacy", and note that its influence has gone well beyond generative approaches. For example, typological and usage-based approaches to language processing typically downplay issues related to the acquisition and evolution of language (e.g., Clark 1996; Hawkins 1994). Similarly, work on language acquisition tends not to consider questions pertaining to the processing and evolution

Morten H. Christiansen. 2016. Language intertwined across multiple timescales: Processing, acquisition and evolution. In N. J. Enfield (ed.), *Dependencies in language*, 55–64. Berlin: Language Science Press. DOI:??

of language (e.g., Cowie 1999; Hirsh-Pasek & Golinkoff 1996; O'Grady 1997), and studies of language evolution usually pay little attention to research on language acquisition and processing (e.g., Botha 2003; Burling 2005; Corballis 2002; Dunbar 1998; Lieberman 2000). In contrast, Christiansen & Chater (2016a) argue that there are strong theoretical constraints between the processing, acquisition and evolution of language—allowing each to shed light on the others—and that key questions within each area can only be fully addressed through an integrated approach. As an example, I briefly discuss how the immediacy of language processing has implications for both language acquisition and evolution.

1 The Now-or-Never Bottleneck

Language happens in the here-and-now. Our memory for acoustic information is incredibly short-lived, disappearing within less than 100 msec (Remez et al. 2010). At the same time spoken language comes at us at a very rapid rate at about 10-15 phonemes per second (Studdert-Kennedy 1986), with the further complication that our auditory system is only able to keep track of about 10 separate (non-speech) sounds per second (Miller & Taylor 1948). To make matters worse, our ability to keep track of sound sequences is also very limited: we are able to recall less than four non-speech sounds (Warren et al. 1969) and only four-to-seven unrelated linguistic items (Cowan 2000; Miller 1956). Thus, during a normal conversation, we are faced with an immense challenge by the combined effects of poor acoustic memory, fast input, and severely limited sequence memory.[1] As a consequence of this *Now-or-Never bottleneck* (Christiansen & Chater 2016b), new material will constantly overwrite and interfere with previous material unless it is processed immediately.

The Now-or-Never bottleneck has direct implications for language processing. To deal with the immediacy of language, Christiansen & Chater (2016b) suggest that the language system must engage in *Chunk-and-Pass* processing: compress and recode language input as rapidly as possible into increasingly more abstract levels of linguistic representation, from sound-based units to words (or word combinations) to discourse-level representations. This passing up of chunks allows for increasingly longer retention of linguistic information at higher levels of linguistic abstraction, consistent with recent neuroimaging data (e.g., Ding et al. 2016; Stephens, Honey & Hasson 2013).

[1] Communication using sign language involves a similar problem (see Christiansen & Chater 2016b for discussion)

The time-sensitive nature of Chunk-and-Pass processing leads to a strong pressure toward incremental processing because chunking will primarily happen across neighboring units, resulting in a bias toward local dependencies (in line with evidence for garden path effects in language comprehension; e.g., Bever 1970). The multiple levels of linguistic structure that result from the Chunk-and-Pass process provides a possible processing-based explanation for why linguistic theories tend to be couched in terms of multiple levels of representation, from phonology and morphology to syntax and discourse.[2] Importantly, though, in the proposed framework, higher levels of representations will contain less of the original detail of the input as it becomes more compressed through repeated Chunk-and-Pass processing.

Because the Now-or-Never bottleneck prevents any significant backtracking, the language system employs prediction to use as much available information as possible to be right first time. In doing so, the processing system will build the most abstract and complete representation that is justified, given the linguistic input—a "good-enough" representation (Ferreira, Bailey & Ferraro 2002; Ferreira & Patson 2007). Through prediction, top-down information from discourse expectations, world knowledge, and so on, is used to guide the incremental interpretation of linguistic input. Language production follows the same principles but in the opposite direction, from discourse representations of the intended message and intonational phrases to words and articulatory motor commands (see Chater & Christiansen 2016; Chater, McCauley & Christiansen 2016 for discussion).

The effects of the Now-and-Never bottleneck go beyond the timescale of processing to the timescale of acquisition. In order to become a competent language user, the child must learn how to create and integrate the right chunks as rapidly as possible, before the input is gone. From this perspective, language acquisition does *not* consist in identifying the right grammar but rather, *language acquisition is learning to process*, to become more efficient at Chunk-and-Pass processing. That is, the child is not a "mini-linguist" but a developing language user, acquiring the necessary skills to comprehend and produce language. To deal with the Now-or-Never bottleneck, the child must learn in the "here-and-now," relying only on currently available information, instead of abstracting over large swaths of data[3]. Learning is therefore local and piecemeal, constrained by lim-

[2] Although this perspective is consistent with standard levels of linguistic abstraction, from phonology through syntax to pragmatics, a complete model might incorporate more fine-grained levels that, for example, would distinguish between multiple levels of discourse representation (e.g., as in Enfield 2013).

[3] The Now-or-Never bottleneck thus has important implications for computational models of

ited memory, in line with item-based approaches to language acquisition (e.g., Tomasello 2003). Children gradually learn to apply top-down knowledge to facilitate Chunk-and-Pass processing via prediction. Thus, predictive abilities emerge over time as children develop their chunking skills and learn to rapidly apply the multiple constraints that are crucial to adult incremental processing (Borovsky, Elman & Fernald 2012).

The theoretical impact of the Now-or-Never bottleneck not only affects the timescales of processing and acquisition, but also extends to the longer timescales of language evolution and change. Given the hypothesis that language evolution may be explained primarily by the cultural evolution of linguistic structure rather than biological adaptations for language (e.g., Christiansen & Chater 2008; Hurford 1999; Smith & Kirby 2008; for a review, see Dediu et al. 2013), we might expect that linguistic patterns that can be processed through the bottleneck will tend to proliferate. That is, language is a product of piecemeal tinkering, with the long-term evolution of language resulting from the compounding of a myriad local short-term processes of language change. This means that *language change is item-based* in nature, with specific changes arising from constraints on Chunk-and-Pass processing—both within and across individuals—providing a possible cognitive foundation for grammaticalization.

The Now-or-Never bottleneck provides a constant pressure towards reduction and erosion across different levels of linguistic representation, from discourse syntacticization and semantic bleaching to morphological reduction and phonetic erosion (see Christiansen & Chater 2016b for further discussion). Language change, more broadly, will be local at the level of individual chunks, consistent with theories of lexical diffusion suggesting that sound change originates in a small set of words and then spreads throughout the vocabulary (e.g., Wang 1977). Similarly, morpho-syntactic change is also predicted to be local in nature, resulting in what Christiansen & Chater (2016b) term "constructional diffusion."

Importantly, the process of piecemeal tinkering that drives item-based language change is subject to constraints deriving not only from Chunk-and-Pass processing but also from the specific trajectory of cultural evolution that a language follows. More generally, in this perspective, there is no sharp distinction

language, many of which use so-called batch-learning either over large corpora (e.g., Perfors, Tenenbaum & Wonnacott 2010) or large memory windows (e.g., Kolodny, Lotem & Edelman 2015) incompatible with psychological constraints on memory. In contrast, the Chunk-Based Learner (McCauley & Christiansen 2014; 2016) was developed with the Now-and-Never bottleneck in mind, providing a computational account of aspects of early language acquisition, including the interconnected nature of comprehension and production (Chater, McCauley & Christiansen 2016).

between language evolution and language change: language evolution is simply the result of language change writ large (see also Heine & Kuteva 2007), constrained by processing and acquisition (see Christiansen & Chater 2016a for more details).

2 Language Intertwined across Multiple Timescales

In this chapter, I have discussed how the Now-or-Never bottleneck not only provides constraints on the processing of language but also on the nature of language acquisition and evolution (with further implications for structure of language itself, as discussed in Christiansen & Chater (2016a,b). Figure 1 provides an illustration of how the Now-or-Never bottleneck affects language across these different timescales.

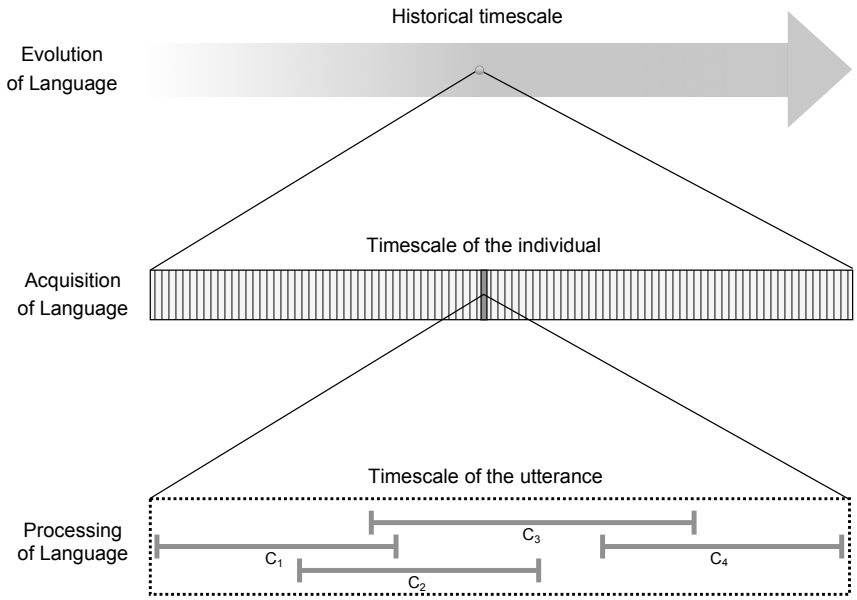

Figure 1: Illustration of how Chunk-and-Pass processing at the utterance level (with the $C_{1\text{-}4}$ referring to different chunks) constrains the acquisition of language by the individual, which, in turn, influences how language evolves through learning and use by groups of individuals on a historical timescale. Adapted from Christiansen & Chater (2016a).

At the timescale of the utterance (seconds), Chunk-and-Pass processing carves the input—or output—into chunks at various levels of linguistic abstraction. At the timescale of the individual (tens of years), these chunks provide the comprehension and production events from which children learn (and adults update) their ability to process language. And, on a historical timescale (hundreds or thousands of years), each learner is part of a community of language users that together change language, based on patterns that are easy to acquire and process. Of course, the Now-or-Never bottleneck works together with other constraints deriving from the brain and body to shape the cultural evolution of language (Christiansen & Chater 2008; Christiansen & Chater 2016a), where the brain and body are embedded in a social network of interactions. Thus, to reach a complete understanding of how language works, we need to study it as intertwined across the multiple timescales of processing, acquisition and evolution.

Acknowledgments

Thanks to Erin Isbilen for comments on a previous version of this chapter.

References

Bever, T. G. 1970. The cognitive basis for linguistic structures. In J. R. Hayes (ed.), *Cognition and the development of language (pp*, 279–362. New York: Wiley & Sons.

Bock, J. K. 1986. Meaning, sound, and syntax: lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12. 575–586.

Borovsky, A., J. L. Elman & A Fernald. 2012. Knowing a lot for one's age: vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology* 112. 417–436.

Botha, R. P. 2003. *Unravelling the evolution of language*. Amsterdam: Elsevier.

Burling, R. 2005. *The talking ape: how language evolved*. Oxford: Oxford University Press.

Chater, N. & M. H Christiansen. 2016. Squeezing through the now-or-never bottleneck: reconnecting language processing, acquisition, change and structure. *Behavioral Brain Sciences* 39.

Chater, N., S.M McCauley & M. H Christiansen. 2016. Language as skill: intertwining comprehension and production. journal of memory and language. *Journal of Memory and Language* 89.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1975. *Reflections on language*. New York: Pantheon Books.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22.

Christiansen, M. H. & N Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31. 489–558.

Christiansen, M.H. & N Chater. 2016a. *Creating language: integrating evolution, acquisition, and processing*. Cambridge, MA: MIT Press.

Christiansen, M.H. & N Chater. 2016b. The now-or-never bottleneck: a fundamental constraint on language. *Behavioral Brain Sciences* 39.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Corballis, M. C. 2002. *From hand to mouth: The origins of language*. Princeton: Princeton University Press.

Cowan, N. 2000. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24. 87–114.

Cowie, F. 1999. *What's within? nativism reconsidered*. New York: Oxford University Press.

Dediu, D., M. Cysouw, S.C. Levinson, A. Baronchelli, M.H. Christiansen, W. Croft, N. Evans, S. Garrod, R. Gray, A. Kandler & E Lieven. 2013. Cultural evolution of language. In P.J. Richerson & M. H. Christiansen (eds.), *Cultural evolution: society, technology, language and religion (pp*, 303–332. Cambridge, MA: MIT Press.

Ding, N., L. Melloni, H. Zhang, X. Tian & D Poeppel. 2016. Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience* 19. 158–164.

Dunbar, R. I. M. 1998. *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.

Enfield, N. J. 2013. *Relationship thinking: agency, enchrony, and human sociality*. New York: Oxford university Press.

Ferreira, F., K. G. Bailey & V Ferraro. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11. 11–15.

Ferreira, F. & N. D Patson. 2007. The "good enough" approach to language comprehension. *Language and Linguistics Compass* 1. 71–83.

Greenberg, J. H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (ed.), *Universals of language (second edition) (pp*, 73–113. Cambridge, MA: MIT Press.

Hawkins, J. A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: a reconstruction*. Oxford: Oxford University Press.
Hirsh-Pasek, K. & R Golinkoff (eds.). 1996. *Action meets word: how children learn verbs*. New York: Oxford University Press.
Hurford, J. 1999. The evolution of language and languages. In R. Dunbar, C. Knight & C. Power (eds.), *The evolution of culture (pp*, 173–193. Edinburgh, UK: Edinburgh University Press.
Jackendoff, R. 2002. *Foundations of language: brain, meaning, grammar, evolution*. New York: Oxford University Press.
Kolodny, O., A. Lotem & S Edelman. 2015. Learning a generative probabilistic grammar of experience: a process-level model of language acquisition. *Cognitive Science* 39. 227–267.
Lieberman, P. 2000. *Human language and our reptilian brain*. Cambridge, MA: Harvard University Press.
McCauley, S.M. & M.H Christiansen. 2014. Acquiring formulaic language: a computational model. *Mental Lexicon* 9. 419–436.
McCauley, S.M. & M.H Christiansen. 2016. *Language learning as language use: A cross- linguistic model of child language development*. Manuscript in preparation.
Miller, G. A. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63. 81–97.
Miller, G. A. & W. G Taylor. 1948. The perception of repeated bursts of noise. *Journal of the Acoustical Society of America* 20. 171–182.
O'Grady, W. 1997. *Syntactic development*. Chicago, IL: University of Chicago Press.
Perfors, A., J.B. Tenenbaum & E Wonnacott. 2010. Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language* 37. 607–642.
Pinker, S. 1994. *The language instinct: how the mind creates language*. New York, NY: William Morrow & Company.
Remez, R. E., D. F. Ferro, K. R. Dubowski, J. Meer, R. S. Broder & M. L Davids. 2010. Is desynchrony tolerance adaptable in the perceptual organization of speech? *Attention, Perception & Psychophysics* 72. 2054–2058.
Smith, K. & S Kirby. 2008. Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B* 363. 3591–3603.

Stephens, G. J., C. J. Honey & U Hasson. 2013. A place for time: the spatiotemporal structure of neural dynamics during natural audition. *Journal of Neurophysiology* 110. 2019–2026.

Studdert-Kennedy, M. 1986. Some developments in research on language behavior. In N. J. Smelser & D. R. Gerstein (eds.), *Behavioral and social science: fifty years of discovery: in commemoration of the fiftieth anniversary of the "ogburn report: recent social trends in the united states" (pp*, 208–248. Washington, DC: National Academy Press.

Tomasello, M. 2003. *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Wang, W.S.-Y (ed.). 1977. *The lexicon in phonological change*. The Hague: Mouton.

Warren, R. M., C. J. Obusek, R. M. Farmer & R. P Warren. 1969. Auditory sequence: confusion of patterns other than speech or music. *Science* 164. 586–587.