特集論文

Linguistic Knowledge by Descent:
An Evolutionary Approach to Stress Typology

Bridget D. Samuels*, Pedro T. Martins** and Cedric Boeckx**, ***

言語知識の由来
—強勢類型への進化的アプローチ—

SUMMARY: We consider here several properties of phonological stress systems, including the *midpoint pathology*, an unattested pattern in which stress is confined to a word-medial syllable in short words but reverts to an edge-based window in longer words. Previous attempts have been made to rule out midpoint systems by eliminating the phonological constraints that yield them, or by alluding to difficulties in learning them. We suggest that a preference for representing word edges in memory and limits on subitization—evolutionarily older "fossil" abilities which are neither specific to humans nor to language—are sufficient to rule out the midpoint pathology. We take the same approach to motivate accentual window size and some left-right asymmetries observed in the typology of attested stress systems. This approach highlights the relevance of descent in accounting for human cognition, as well as the benefits that evolutionary thinking can bring to the study of language.

Key words: phonology, stress, midpoint pathology, evolution, typology

1.　Introduction

　One of the chief goals of modern linguistics is to discover the limits of cross-linguistic variation and describe the hypothesis space available to a language learner. An immediate difficulty that arises in the pursuit of this goal is that there are multiple reasons why a certain pattern may be unattested. Some may be diachronically impossible: they could in principle be computed by the language faculty, but never arise through natural historical processes; others may not be learnable or computable (Hale and Reiss 2000, 2008). It is not always obvious which of these factors is at play. Moreover, there may not be a sharp divide between, for example, learnable and non-learnable. Learnability may be influenced by a number of biases that humans bring to the process of language acquisition, or by phonetic factors that make certain contrasts more or less difficult to hear/produce. These biases can produce "soft" constraints, which may explain why there are few absolutely and "non-accidentally" exceptionless universals in phonology (Blevins 2004, Kiparsky 2006). Some approaches to phonology seek to encode these biases in the grammar, which is often called "phonetic grounding" (see, for example, the contributions in Hayes et al. 2004). Other approaches agree that such biases are indeed of great value for explaining phonological typology but contend that they need not (or more strongly, should not) be recapitulated in the formal grammar, for doing so creates undesirable redundancy (Blevins 2004, Hale and Reiss 2008, Samuels 2011). We take the latter approach, suggesting here that extralinguistic biases and properties of our cognition may be sufficient to explain some aspects of the observed cross-linguistic variation in phonological stress systems. In particular, we consider several cross-linguistic typological generalizations about phonological stress systems, including the case of the "midpoint pathology" (Hyde 2008), as well as accentual window size and left-right asymmetries observed in stress placement.

2.　The Midpoint Pathology

　Briefly, the midpoint pathology describes a type of unattested system in which stress is confined to a word-medial syllable in short words but reverts to an

---

　*　University of Southern California（南カリフォルニア大学）
　**　Universitat de Barcelona（バルセロナ大学）
***　Catalan Institute for Research and Advanced Studies（ICREA）

edge-based window in longer words. An earlier account (Kager 2012) attempts to rule out such systems computationally, by eliminating the constraints that could generate it from CON, the set of universal, ranked constraints that are held to generate all possible languages according to Optimality Theory (OT) (Prince and Smolensky 1993 [2004]). Let's call this Humboldt's solution, since this is an answer that is formulated at the level of investigation that Chomsky (1986) refers to as "Humboldt's problem": identifying the right set of computational properties that make up linguistic knowledge. More recently, Stanton (2016) has instead proposed that midpoint systems are unattested because they represent a learning problem. Let's call this Plato's solution, an allusion to "Plato's problem", or the logical problem of language acquisition identified by Chomsky (1986). We contend that a preference for representing word edges in memory and constraints on the number of objects that can be represented in working memory, both of which are part of our primate cognitive inheritance, are sufficient to rule out the midpoint pathology. Let's call this Darwin's solution, since it addresses "Darwin's problem", or the problem of language evolution (Boeckx 2011). If correct, our argument illustrates how properties of cognition that pre-date human language, termed "fossils" in this volume, may be seen to shape phonology.

The types of languages relevant to the present discussion are those with metrical window systems (Kager 2012)[1]. In these languages, stress falls within a two- or three-syllable window from the left or right edge of a word. *ExtendedLapse (*ExtLapse) constraints penalizing sequences of three stressless syllables at a word edge and *Lapse constraints penalizing sequences of two stressless syllables at a word edge, as in (1), have been proposed to enforce these windows (Elenbaas and Kager 1999, Gordon 2002)[2].

(1) Optimality theoretical constraints that penalize windows of stressable syllables at word edges
 a. *ExtLapseL: Assign one * if the initial three syllables of the word are stressless.
 b. *ExtLapseR: Assign one * if the final three syllables of the word are stressless.
 c. *LapseL: Assign one * if the initial two syllables of the word are stressless.
 d. *LapseR: Assign one * if the final two syllables of the word are stressless.

Midpoint pathologies arise when two of these contextualized *Lapse or *ExtLapse constraints are both highly ranked and therefore compete. One such ranking, adapted from Stanton (2016), is illustrated in (2) with strings ranging from one to seven syllables in length. The window in which a stressed syllable would satisfy *ExtLapseL is shown in curly brackets and the window in which a stressed syllable would satisfy *ExtLapseR is shown in square brackets. The potentially stressable syllables are underlined.

(2) Constraint ranking resulting in midpoint pathology (Stanton 2016).
 *ExtLapseL >> *ExtLapseR
 a. [{$\underline{\sigma}$}]
 b. [{$\sigma\,\underline{\sigma}$}]
 c. [{$\underline{\sigma}\,\underline{\sigma}\,\underline{\sigma}$}]
 d. {$\sigma[\underline{\sigma}\,\underline{\sigma}]\,\sigma$}
 e. {$\sigma\,\sigma[\underline{\sigma}]\,\sigma\,\sigma$}
 f. {$\underline{\sigma}\,\underline{\sigma}\,\underline{\sigma}$}[$\sigma\,\sigma\,\sigma$]
 g. {$\underline{\sigma}\,\underline{\sigma}\,\underline{\sigma}$}$\sigma[\sigma\,\sigma\,\sigma$]

In the hypothetical midpoint system in (2), stress may fall on any syllable in short (<4 syllable) words because all the syllables fall within the window that satisfies both *ExtLapse constraints. In words of four or five syllables, stress is confined to the middle of the word because these windows only partially overlap. Word with six or more syllables must be stressed in the three-syllable window at the leftmost edge of the word because there is no way to satisfy both constraints simultaneously, and *ExtLapseL dominates *ExtLapseR.

Kager (2012) proposes a computational method (Humboldt's solution) of ruling out such unwanted rankings, namely rejecting *(Ext)Lapse constraints in favor of an account that involves weakly layered, maximally ternary feet. We will not consider this approach here, but it should be noted that eliminating this family of constraints appears to preclude a purely grid-based OT evaluation of stress. Thus, for those who pursue an OT analysis but reject a foot-based approach (on the relative merits of grid- and foot-based approaches in OT, see Gordon 2011), there is a strong motivation to (a) determine whether midpoint systems should indeed be ruled out, and if so, (b) find an alternative that can salvage *LAPSE constraints. This is the work undertaken by Stanton (2016).

Stanton (2016) argues that 3.26% of all languages should be midpoint systems based on the factorial typology (the fraction of rankings that would produce such a system out of the factorial of the total number of constraints) if all rankings of the set proposed by Kager (2012) are equally probable. We would therefore expect

to see 16 such languages in the StressTyp database (http://stresstyp.leidenuniv.nl/), which included 510 languages at the time of Stanton's writing, but there are no known languages with stress patterns of this type[3]. Based on a binomial test, the difference between the expected (16) and attested (0) number of midpoint systems in StressTyp is significant ($p<0.001$). Stanton therefore concludes that the absence of midpoint systems deserves an explanation.

While this may appear to be a theory-internal problem, since the midpoint pathology is in some sense "created" by a very particular set of constraints posited in a certain theory of phonology, we emphasize that this is not the case. It is true that other methods of computing stress, such as the ones proposed by Kager (2012) or (to pick a classic non-OT account) Halle and Vergnaud (1987), do not permit/generate midpoint systems. This would be a computation-level explanation, or as we are calling it here, Humboldt's solution. We pursue an explanation from a different angle, as outlined above. In what follows, we suggest that the absence of midpoint systems, maximal window size, and asymmetry between initial and final window systems may be derived from learning biases shared by humans and other species that are independently attested in the literature. Again, we call this Darwin's solution. If complex abilities—including linguistic ones—are decomposed into simpler mechanisms, which can then be found to be shared by domains other than language and species other than humans, their origins can be traced further back in evolutionary history, and we come closer to understanding what, if anything, had to emerge in the human lineage in order to make human phonology possible and display the properties that it does. This is one of the broader goals of our work (Martins and Boeckx 2014, Samuels 2009a, 2009b, 2011, 2012a, 2012b, 2015a, 2015b, Samuels et al. 2016).

3. Edge Preference

Encoding the serial order of linguistic objects is a crucial part of language learning, particularly in phonology and morphology. For example, no language treats [kat], [tak], [atk], [akt], [tka], and [kta] as interchangeable variants of the same lexical item. The nature of serial order encoding has been studied in both infants and adults, as well as in non-human primates.

Research on sequential encoding in human infants has demonstrated repeatedly that stimuli presented first and last have an advantage in long-term memory over those in the middle of a sequence. This has been shown for both non-linguistic stimuli, such as faces and visual patterns (Gulya et al. 1998, 1999, 2001), and for linguistic ones. Even newborns appear to have enhanced encoding of edge syllables (Ferry et al. 2015). Importantly for the present purposes, melody and prosody are helpful and perhaps necessary cues for infants to detect changes in the order of words in the middle of a sequence (Mandel et al. 1996, Thiessen and Saffran 2009). It may be, then, that prosodic contours help to define the edges of a sequence and cement its serial order in memory relative to those edge positions. This would be one of several prosodic bootstrapping processes used by children early in life (Pinker 1984); see Vihman (2014) for an overview. A bidirectional relationship between memory and cues from the linguistic input has been suggested (Benavides-Varela and Mehler 2015): memory mechanisms enable the position and content of edge material to be encoded, while cues such as lexical stress and prosody add salience to certain positions and facilitate the segmentation of the speech stream into chunks, aiding with the memorization of sequence middles. Taking this view, the location of stress near word edges (and boundary tones near phrase edges) is crucial for their accurate positional encoding, which in turn mediates the encoding of other elements further from the edges.

In adults, serial position effects have been amply demonstrated as well (see Benavides-Varela and Mehler 2015 for a review): the first and last items in a list of words or syllables are remembered better than those in the middle. Moreover, it is easier for adults to extract generalizations that occur at a word edge than ones that involve elements in the middle of a word (Endress et al. 2005, Endress and Mehler 2010). Endress et al. (2009b) have argued that the edges of linguistic sequences are "anchors" that help language learners to extract patterns, and mention stress specifically as a phenomenon that is crucially served by edge-based positional regularities. They further contend that this serial encoding mechanism may reflect a more general, phylogenetically older mechanism that was recruited for language. This is so because the two perceptual/ memory processes they point to—one that computes identity relations (see Samuels 2014) and one that computes position codes in edges—are not linguistic in nature, even though they are crucial to human language. A "U-shaped" or "bowed" serial position function, with improved memory of the first items in a list (primacy effect) as well as the last (recency effect), is typical of both linguistic and non-linguistic performance in multiple modalities; recency and primacy effects have

also been demonstrated in rhesus monkeys, capuchin monkeys, and pigeons (Wright 2007). Benavides-Varela and Mehler (2015, p. 220) conclude that edge-based regularities may "appeal to a positional memory mechanism that is not specific to the human species or to the linguistic knowledge, yet it appears to influence humans' linguistic representations early in life."

Studies on sequence learning in non-human primates and in birds support this position (see Endress et al. 2009a, Dehaene et al. 2015). Based on their performance on several behavioral tasks, it can be concluded that macaques are able to represent the order of list items in memory. For example, they are able to generalize from pairwise orderings (a, b), (b, c), and (c, d) to infer an overall order of (a, b, c, d) (Merritt and Terrace 2011). Temporal ordering of motor sequences is also crucial to a number of other abilities across the animal kingdom (Endress et al. 2009a), such as foraging routines (Greenfield 1991, 1998, Fujita 2009) and song structure in birds and whales (Payne 2000, Doupe and Kuhl 1999). Neurons in the dorsolateral prefrontal and intraparietal cortex have been shown to tune to the order of items, responding identically to items in a particular ordinal position regardless of the identity of the item (Ninokura et al. 2004, Nieder et al. 2006, Nieder 2012). Chimpanzees are also stronger at representing sequence edges and therefore more sensitive to generalizations that involve positional regularities involving edges (Endress et al. 2009a)[4]. Given that these species do not have language, it is likely that they share with us a more generic edge-based memory mechanism that humans inherited and recruited for language.

Taken together, these studies support the conclusion that the ability to encode ordinal knowledge is neither unique to humans nor to language, but could be exploited in language for the encoding of stress, among other things. The goal of the present paper is not to argue for one method of stress computation over another. However, it should again be noted that, as far as we are aware, the midpoint pathology is only predicted in schemes that "pull" stress simultaneously to both edges of the word (Kager 2012); these opposing forces are provided by the set of symmetric *(Ext)Lapse constraints in Figure (1).

4. Subitization, Prosodic Bootstrapping, and Asymmetry

While the discussion of edge preference and positional memory above goes some distance towards explaining why stress is assigned with respect to the edges of words, it does not explain why metrical windows typically encompass the edgemost two or three syllables. We suggest that limits on the size of a metrical window may pertain to the number of visual or tactile objects that can be subitized (Kaufman et al. 1949).

Literature on the number of items that can be subitized (Kaufman et al. 1949)—apprehended without counting—suggests that numerosities up to 3 can be perceived rapidly, accurately, and confidently. Beyond this range, the "cost" (in terms of effort, speed, accuracy, and confidence) rises dramatically with each additional item presented (Trick and Pylyshyn 1994). This phenomenon was initially discovered in the visual modality, and similar findings have since been reported for tactile (Riggs et al. 2006, Plaisier et al. 2009) and auditory stimuli (Camos and Tillmann 2008, Ruusuvirta and Astikainen 2016). Similar constraints have been revealed for both humans and other species including rhesus and capuchin monkeys (Beran et al. 2011), chimpanzees (Tomonaga and Matsuzawa 2002), and many other more distantly related vertebrates, including dogs and several varieties of fish, leading to the conclusion that the "object-file system" underlying this ability was already in place prior to the divergence of bony fish and tetrapods (see Dehaene 1997, Agrillo et al. 2012 and references therein). If syllable-sized chunks are the appropriate unit to consider for the purposes of stress computation, then this evolutionarily ancient system may explain the observed maximal size of stress windows. It may also relate to the size of feet (in frameworks that admit them), which can include two or three syllables; tone assignment, which may require counting up to four moras (Marlo et al. 2015); and infixation and reduplication processes, which may anchor to the first or second element (consonant, vowel, segment, syllable, or foot) from either edge of a word (Samuels 2010). Chunks of phonological material comprising similar numbers of elements are relevant at different levels of computation, overlain on one another: segments per syllable, syllables per foot, and so forth.

Looking at 160 window systems represented in StressTyp, shown in Table 1, it also appears that there

Table 1   Summary of stress window systems in StressTyp (Kager 2012, p. 1464)

	Phonological	Lexical	Mixed	Total
Final two σs	59	20	3	82
Final three σs	22	12	4	38
Initial two σs	26	10	3	39
Initial three σs	1	0	0	1

is a preference for the stressable window to gravitate towards the right edge of a word rather than the left, as already noted by Hyman (1977).

As Kager (2012) mentions, there are a handful of other known systems with initial three-syllable windows not included in StressTyp; a small number of languages have also been analyzed as having final four- or five-syllable window systems. In these latter cases, which include Kashaya Pomo (Buckley 1994) and some varieties of Arabic (Hayes 1995), stress has been analyzed as falling on the head of the second foot from the right edge under certain circumstances. However, again recall that foot-based analyses are unavailable in system of stress assignment which uses the constraints in (1).

We contend that the asymmetry observed in Table 1 results from the culminative and delimitative (also known as demarcative) functions of stress (Trubetzkoy 1939). The culminative function of stress reflects the fact that (primary) stress must occur precisely once per word. Stress can therefore serve as an important cue for word segmentation, which is its delimitative or demarcative function. Upon hearing a sequence of two stresses, a listener can infer that a word boundary occurred somewhere between them (Gambell and Yang 2005). The sensitivity of infants to the location of stress suggests this strategy may indeed be used during first language acquisition (Jusczyk et al. 1999). This task of locating word boundaries may be facilitated if stress occurs in a predictable place relative to the boundary (Hyman 1977, Cutler and Norris 1988). This is another type of prosodic bootstrapping; see Gambell and Yang (2005) for a review of the literature on this "Metrical Segmentation Strategy". If stress is anchored to the right edge of the word, a listener knows when to expect the word boundary upon hearing a stressed syllable. On the other hand, if stress is anchored to the left edge of the word, upon hearing a stressed syllable, the listener is simply informed that the word boundary has already passed and must backtrack to determine where the boundary should have been. Thus, stress systems with the window at the right edge may better facilitate parsing[5].

The next logical question to ask is why stress falls more frequently on the penult or antepenult than on the ultima in stress windows that fall on the right edge of the word. Notably, this pattern is asymmetric: initial stress is more common than peninitial (second syllable from the left), and postpeninitial stress is very rare. In foot-based analyses, the generalization here is that trochaic feet are more frequent than iambic ones at word edges (Halle and Vergnaud 1987, Hayes 1995). Especially in analyses that cannot appeal to feet, this

still begs the question of why final stress is underrepresented—or, why are final syllables often extrametrical—in light of what we have said so far about the edge preference. Building further on this observation, it has been noted that many languages shift stress from the final syllable to the penult, but none shift stress from the initial syllable to the peninitial one (Hyman 1977, Karvonen 2008). This avoidance of final stress prevents perceptual crowding with higher-level prosodic boundary tones, which tend to mark the right edge of a phrase rather than the left (Gordon 2000, Karvonen 2008). For example, a high tone on a stressed syllable of a word (H*) and a low boundary tone on the final syllable in an intonational phrase (L%) would be crowded on the same syllable in a language with final stress.

Interestingly, a purely perceptual mechanism may be available to explain stress shift in cases like these (Shattuck-Hufnagel 1995, Samuels 2009a). One prominent example is the so-called Rhythm Rule in English, which describes alternations like the one between the isolation or counting form of the numeral *thìrtéen* and the form found in a phrase like *thírtèen mén*. One analysis consistent with this perceptual account posits that both syllables of *thirteen* bear lexical stress equally, but primary stress is perceived on the final syllable when it is pronounced in isolation as the result of phrasal stress or a phrase-final pitch accent (typically L% in a declarative context). The perceived stress shift in *thirteen men* then arises not from a phonological process moving or deleting the stress on the final syllable of *thirteen*, but from the interaction of the lexical stress with a phrasal pitch accent (assigned to the leftmost stress in the highest element in the complement domain of the D(eterminer); see Samuels 2009a) that docks on the first stressable unit of the phrase, in this case the first syllable of *thirteen*. Similar analyses may explain processes in other languages that are claimed to shift stress from the final to penultimate syllable, as noted above.

Collectively, the studies presented in this section suggest that a combination of subitization, parsing considerations, and the interaction of stress with boundary tones may explain the maximal size of the stressable windows seen cross-linguistically as well as asymmetries in the typology of stress window systems.

5. Conclusions

Unattested systems pose a challenge for any theory that aims to account for the limits of linguistic variation and acquisition. A theory must walk a fine line between being too restrictive, thereby ruling out attested

systems, and being too permissive, thereby predicting unattested systems. Furthermore, it may be difficult to assess whether a theory is too permissive, because it is not always clear why a certain pattern is unattested. The present work provides an account of why midpoint stress systems may be unattested and why certain asymmetries may be observed in stress systems, based on principles that are phonologically theory-neutral. Our intention here is not to argue for a particular account of stress computation, but to uncover some extralinguistic biases that may help to explain some aspects of the cross-linguistic typology, and which therefore—in our view, though it may not be the mainstream in phonological theory—need not be duplicated via constraints in the grammar.

By pointing to discoveries and insights from outside of linguistics, namely "fossil" abilities we share with our primate relatives and which we also use in domains other than language, we address Darwin's problem in the context of some aspects of phonological stress systems by proposing an account that benefits from independent support, consistent with evolutionary thinking. This approach is consistent with other accounts that propose primitives of learning language along these lines (Gervain and Mehler 2010). We use the properties of phonological stress systems discussed here not only as an example of how a linguistic phenomenon can be explained by alluding to research outside of linguistics, but also more generally as a way to advocate for an interdisciplinary approach to the study of language that integrates different levels of explanation and fields of expertise, which we believe will be most fruitful.

Notes

1) Not all languages are of this type. For example, in languages such as Kashmiri (Morén 2000) and Nanti (Crowhurst and Michael 2005), heavy syllables attract stress regardless of their position in a word (often called weight-to-stress principle effects). Furthermore, it should be noted that morphological structure plays a role in confining the position of stress in polymorphemic words (see Samuels 2011 for an analysis in terms of word-internal phase domains). We set these issues aside here.

2) We follow standard conventions in which assigning a * signals a constraint violation, σ denotes a syllable, and A $>>$ B indicates that Constraint A dominates (i.e., is ranked higher than) Constraint B.

3) In a few languages, the possible locations of stress depend on word length. These include North Kyungsan Korean (Kenstowicz and Sohn 2001) and Içuã Tupi (Abrahamson 1968). For example, in North Kyungsan

Korean, stress may fall on the penultimate or final syllable in words of three syllables or fewer, but falls exclusively on the penult in words of four or more syllables. Such systems are underattested based on the factorial typology assuming the constraints described in the main text. See Stanton (2016, §5.2) for further discussion.

4) There is substantial evidence from studies since Kuhl and Miller (1975) that many perceptual abilities traditionally thought to be specific to humans and their speech, namely categorical perception of speech sounds, perception of prototypical vowels and compensation for coarticulated phonemes, are in fact shared by several other species (see, *inter alia*, Patterson and Pepperberg 1994, 1998, Endress et al. 2009a, Samuels 2009).

5) This may seem to run counter to the findings of Cutler and Butterfield (1992), who found that native English speakers tend to posit a boundary prior to a stressed syllable. This reflects a correct generalization about the English lexicon, which has a predominance of words with strong initial syllables (see, e.g., Cutler and Carter 1987). Such a bias may therefore arise from experience rather than an innate tendency. We claim only that parsing may be facilitated in systems with stress anchored to the right rather than the left, though both types of languages are clearly learnable.

Acknowledgements

References

Abrahamson, A. (1968) "Contrastive distribution of phoneme classes in Içuã Tupi." *Anthropological Linguistics* 10, 11–21.

Agrillo, C., L. Piffer, A. Bisazza and B. Butterworth (2012) "Evidence for two numerical systems that are similar in humans and guppies." *PLoS ONE* 7, e31923.

Benavides-Varela, S. and J. Mehler (2015) "Verbal positional memory in 7-month-olds." *Child Development* 86, 209–223.

Beran, M. J., S. Decker, A. Schwartz and N. Schultz (2011) "Monkeys (*Macaca mulatta* and *Cebus apella*) and human adults and children (*Homo sapiens*) compare subsets of moving stimuli based on numerosity." *Frontiers in Psychology* 2, 61.

Blevins, J. (2004) *Evolutionary phonology*. Cambridge: Cambridge University Press.

Boeckx, C. (2011) "Some reflections on Darwin's problem in the context of Cartesian biolinguistics." In A. M. Di Sciullo and C. Boeckx (eds.) *The biolinguistic enterprise*, 42–64. Oxford: Oxford University Press.

Buckley, E. (1994) "Persistent and cumulative extrametricality in Kashaya." *Natural Language and Linguistic*

Theory 12, 423–464.

Camos, V. and B. Tillmann (2008) "Discontinuity in the enumeration of sequentially presented auditory and visual stimuli." *Cognition* 107, 1135–1143.

Chomsky, N. (1986) *Knowledge of language*. New York: Praeger.

Crowhurst, M. J. and L. D. Michael (2005) "Iterative footing and prominence-driven stress in Nanti (Kampa)." *Language* 81, 47–95.

Cutler, A. and S. Butterfield (1992) "Rhythmic cues to speech segmentation: Evidence from juncture misperception." *Journal of Memory and Language* 31, 218–236.

Cutler, A. and D. Carter (1987) "The predominance of strong initial syllables in the English vocabulary." *Computer Speech and Language* 2, 133–142.

Cutler, A. and D. Norris (1988) "The role of strong syllables in segmentation for lexical access." *Journal of Experimental Psychology: Human Perception and Performance* 14, 113–121.

Dehaene, S. (1997) *The number sense: How the mind creates mathematics*. Oxford: Oxford University Press.

Dehaene, S., F. Meyniel, C. Wacongne, L. Wang and C. Pallier (2015) "The neural representation of sequences: From transition probabilities to algebraic patterns and linguistic trees." *Neuron* 88(1), 2–19.

Doupe, A. J. and P. K. Kuhl (1999) "Birdsong and human speech: Common themes and mechanisms." *Annual Review of Neuroscience* 22, 567–631.

Elenbaas, N. and R. Kager (1999) "Ternary rhythm and the lapse constraint." *Phonology* 16, 273–329.

Endress, A. D., S. Carden, E. Versace and M. D. Hauser (2009a) "The apes' edge: Positional learning in chimpanzees and humans." *Animal Cognition* 13, 483–495.

Endress, A. D. and J. Mehler (2010) "Perceptual constraints in phonotactic learning." *Journal of Experimental Psychology: Human Perception and Performance* 36, 235–250.

Endress, A. D., M. Nespor and J. Mehler (2009b) "Perceptual and memory constraints on language acquisition." *Trends in Cognitive Sciences* 13, 348–353.

Endress, A. D., B. Scholl and J. Mehler (2005) "The role of salience in the extraction of algebraic rules." *Journal of Experimental Psychology: General* 134, 406–419.

Ferry, A. L., A. Fló, P. Brusini, L. Cattarossi, F. Macagno, M. Nespor and J. Mehler (2015) "On the edge of language acquisition: Inherent constraints on encoding multisyllabic sequences in the neonate brain." *Developmental Science* 19(3), 488–503.

Fujita, K. (2009) "A prospect for evolutionary adequacy: Merge and the evolution and development of human language." *Biolinguistics* 3, 138–153.

Gambell, T. and C. Yang (2005) "Word segmentation: Quick but not dirty." Ms., Yale University.

Gervain, J. and J. Mehler (2010) "Speech perception and language acquisition in the first year of life." *Annual Review of Psychology* 61, 191–218.

Gordon, M. (2000) "The tonal basis of final weight criteria." *Proceedings of the Chicago Linguistics Society* 36, 141–156.

Gordon, M. (2002) "A factorial typology of quantity insensitive stress." *Natural Language and Linguistic Theory* 20, 491–552.

Gordon, M. (2011) "Metrical stress theory." In J. A. Goldsmith, J. Riggle and A. C. L. Yu (eds.) *The handbook of phonological theory*, 2nd edition, 141–163. Chichester: Wiley-Blackwell.

Greenfield, P. M. (1991) "Language, tools, and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior." *Behavioral and Brain Sciences* 14, 531–595.

Greenfield, P. M. (1998) "Language, tools, and brain revisited." *Behavioral and Brain Sciences* 21, 159–163.

Gulya, M., L. Galluccio, A. Wilk and C. Rovee-Collier (2001) "Infants' long-term memory for a serial list: Recognition and reactivation." *Developmental Psychobiology* 38, 174–185.

Gulya, M., C. Rovee-Collier, L. Galluccio and A. Wilk (1998) "Memory processing of a serial list by young infants." *Psychological Science* 9, 303–307.

Gulya, M., B. Sweeney and C. Rovee-Collier (1999) "Infants' memory processing of a serial list: List Length effects." *Journal of Experimental Child Psychology* 73, 72–91.

Hale, M. and C. Reiss (2000) "Phonology as cognition." In N. Burton-Roberts, P. Carr and G. Docherty (eds.) *Phonological knowledge: Conceptual and empirical issues*, 161–184. Oxford: Oxford University Press.

Hale, M. and C. Reiss (2008) *The phonological enterprise.* Oxford: Oxford University Press.

Halle, M. and J.-R. Vergnaud (1987) *An essay on stress*. Cambridge, MA: MIT Press.

Hayes, B. (1995) *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Hayes, B., R. Kirchner and D. Steriade, eds. (2004) *Phonetically-based phonology*. Cambridge: Cambridge University Press.

Hyde, B. (2008) "Alignment continued: Distance-sensitivity, order-sensitivity, and the midpoint pathology." Ms., Washington University in St. Louis. ROA-998.

Hyman, L. M. (1977) "On the nature of linguistic stress." In L. M. Hyman (ed.) *Studies in stress and accent*, 37–82. SCOPIL #4. Los Angeles: University of Southern California.

Jusczyk, P. W., D. Houston and M. Newsome (1999) "The beginnings of word segmentation in English-learning infants." *Cognitive Psychology* 39, 159–207.

Kager, R. (2012) "Stress in windows: Language typology and factorial typology." *Lingua* 122, 1454–1493.

Karvonen, D. (2008) "Explaining nonfinality: Evidence from Finnish." In C. B. Chang and H. J. Haynie (eds.) *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 306–314. Somerville, MA: Cascadilla Proceedings Project.

Kaufman, E., M. Lord, T. Reese and J. Volkmann (1949) "The discrimination of visual number." *American Journal of Psychology* 62, 498–525.

Kenstowicz, M. and H.-S. Sohn (2001) "Accentual adapta-

tions in North Kyungsan Korean." In M. Kenstowicz (ed.) *Ken Hale: A life in language*, 239–270. Cambridge, MA: MIT Press.

Kiparsky, P. (2006) "The amphichronic program vs. Evolutionary Phonology." *Theoretical Linguistics* 32, 217–236.

Kuhl, P. K. and J. D. Miller (1975) "Speech perception by the chinchilla." *Science* 190, 69–72.

Mandel, D., D. Kemler-Nelson and P. W. Jusczyk (1996) "Infants remember the order of words in a spoken sentence." *Cognitive Development* 11, 181–196.

Marlo, M., L. C. Mwita and M. Paster (2015) "Problems in Kuria H tone assignment." *Natural Language and Linguistic Theory* 33, 251–265.

Martins, P. T. and C. Boeckx (2014) "Attention mechanisms and the mosaic evolution of speech." *Frontiers in Psychology: Language Sciences* 5, 1463.

Merritt, D. and H. Terrace (2011) "Mechanisms of inferential order judgments in humans (*Homo sapiens*) and rhesus monkeys (*Macaca mulatta*)." *Journal of Comparative Psychology* 125, 227–238.

Morén, B. (2000) "The puzzle of Kashmiri stress: Implications for weight theory." *Phonology* 17, 365–396.

Nieder, A. (2012) "Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices." *Proceedings of the National Academy of Sciences of the USA* 109, 11860–11865.

Nieder, A., I. Diester and O. Tudusciuc (2006) "Temporal and spatial enumeration processes in the primate parietal cortex." *Science* 313, 1431–1435.

Ninokura, Y., H. Mushiake and J. Tanji (2004) "Integration of temporal order and object information in the monkey lateral prefrontal cortex." *Journal of Neurophysiology* 91, 555–560.

Patterson, D. and I. M. Pepperberg (1994) "A comparative study of human and grey parrot phonation I: Acoustic and articulatory correlates of vowels." *Journal of the Acoustical Society of America* 96, 634–648.

Patterson, D. and I. M. Pepperberg (1998) "A comparative study of human and grey parrot phonation II: Acoustic and articulatory correlates of stop consonants." *Journal of the Acoustical Society of America* 103, 2197–2213.

Payne, K. (2000) "The progressively changing songs of humpback whales: A window on the creative process in a wild animal." In N. Wallin, B. Merker and S. Brown (eds.) *The origins of music*, 135–150. Cambridge, MA: MIT Press.

Pinker, S. (1984) *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Plaisier, M. A., W. M. Bergmann Tiest and A. M. Kappers (2009) "One, two, three, many—subitizing in active touch." *Acta Psychologica* 131, 163–170.

Prince, A. and P. Smolensky (1993 [2004]) *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell.

Riggs, K. J., L. Ferrand, D. Lancelin and A. Simpson (2006) "Subitizing in tactile perception." *Psychological Science* 17, 271–272.

Ruusuvirta, T. and P. Astikainen (2016) "Preattentive and attentive responses to changes in small numerosities of tones in adult humans." *Brain Research* 1634, 68–74.

Samuels, B. D. (2009a) *The structure of phonological theory.* Ph.D. dissertation, Harvard University, Cambridge, MA.

Samuels, B. D. (2009b) "The Third Factor in phonology." *Biolinguistics* 3, 355–382.

Samuels, B. D. (2010) "The topology of infixation and reduplication." *The Linguistic Review* 27, 131–176.

Samuels, B. D. (2011) *Phonological architecture: A biolinguistic perspective.* Oxford: Oxford University Press.

Samuels, B. D. (2012a) "Animal minds and the roots of human language." In C. Boeckx, M. Horno and J. Mendívil (eds.) *Language, from a biological point of view*, 290–313. Newcastle upon Tyne: Cambridge Scholars Publishing.

Samuels, B. D. (2012b) "The emergence of phonological forms." In A. M. Di Sciullo (ed.) *Towards a biolinguistic understanding of grammar. Essays on interfaces*, 193–213. Amsterdam: John Benjamins.

Samuels, B. D. (2014) "On the biological origins of linguistic identity." In H. van Riemsdijk and K. Nasukawa (eds.) *Linguistic identity*, 341–365. Berlin: Mouton de Gruyter.

Samuels, B. D. (2015a) "Biolinguistics in phonology: A prospectus." *Phonological Studies* 18, 161–171.

Samuels, B. D. (2015b) "Can a bird brain do phonology?" *Frontiers in Psychology: Language Sciences* 6, 1082.

Samuels, B. D., M. D. Hauser and C. Boeckx (2016) "Looking for UG in animals: A case study in phonology." In I. Roberts (ed.) *The Oxford handbook of Universal Grammar*, 527–546. Oxford: Oxford University Press.

Shattuck-Hufnagel, S. (1995) "The importance of phonological transcription in empirical approaches to 'stress shift' versus 'early accent': Comments on Grabe and Warren, and Vogel, Bunnell, and Hoskins." In B. Connell and A. Arvaniti (eds.) *Papers in laboratory phonology IV*, 128–140. Cambridge: Cambridge University Press.

Stanton, J. (2016) "Learnability shapes typology: The case of the midpoint pathology." *Language* 92, 753–791.

Thiessen, E. and J. Saffran (2009) "How the melody facilitates the message and vice versa in infant learning and memory." *Annals of the New York Academy of Sciences* 1169, 225–233.

Tomonaga, M. and T. Matsuzawa (2002) "Enumeration of briefly presented items by the chimpanzee (*Pan troglodytes*) and humans (*Homo sapiens*)." *Animal Learning and Behavior* 30, 143–157.

Trick, L. M. and Z. W. Pylyshyn (1994) "Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision." *Psychological Review* 101, 80–102.

Trubetzkoy, N. S. (1939*) Grundzüge der Phonologie*. Prague: Cercle Linguistique du Prague.

Vihman, M. M. (2014) *Phonological development: The first two years*, 2nd edition. Chichester: Wiley-Blackwell.

Wright, A. A. (2007) "An experimental analysis of memory processing." *Journal of the Experimental Analysis of Behavior* 88, 405–433.