Modeling cultural evolution

Language acquisition as multiple-cue integration

Morten H. Christiansen
Department of Psychology, Cornell University, Ithaca, NY, USA,
Santa Fe Institute, NM, USA

Computational and mathematical modeling has revealed that cultural evolution may have played a key role in the evolution of language. In this chapter, I explore the hypothesis that processes of cultural transmission have to a large extent shaped language to fit domain-general constraints deriving from the human brain. An implication of this view is that much of the neural hardware involved in language is not specific to it. But how could language have evolved to be as complex as it is without language-specific constraints? Based on computational modeling of the cultural evolution of language, I propose that language has evolved to rely on a multitude of probabilistic information sources for its acquisition, allowing it to be as expressive as possible while still being learnable by domain-general learning mechanisms. Empirical predictions are derived from this perspective regarding the role of phonological cues in the learning of basic aspects of syntax. These predictions are corroborated by results from corpus analyses, computational modeling, and human experimentation, suggesting that the integration of phonological cues with other types of information is integral to the computational architecture of our language capacity. I conclude by considering how computational modeling of cultural evolution can help us understand the evolution of language.

1. Introduction

The past couple of decades have seen an explosion of research on language evolution, initially fueled by Pinker and Bloom's (1990) groundbreaking article arguing for natural selection of biological structures dedicated to language. With the new millennium, however, a shift has occurred toward explaining language evolution in terms of cultural evolution rather than biological adaptation. Indeed, theoretical and computational considerations indicate that there are substantial restrictions on what linguistic properties can evolve through natural selection (e.g. Chater, Reali & Christiansen 2009;

Christiansen, Reali & Chater 2011). In contrast, cultural evolution is now emerging as the main paradigm for understanding the evolution of language.

Computational modeling has been one of the key factors in bringing about the change in focus from biological to cultural evolutionary processes in theorizing about language evolution (see Brighton, Smith & Kirby 2005, for a review). Although modeling of the biological evolution of language also exists (e.g. Nowak, Komarova & Nyogi 2002), it has been easier to derive empirically testable predictions from the modeling of cultural evolution. For example, Kirby, Dowman and Griffiths (2007) used Bayesian modeling to show that cultural transmission across generations of learners could turn weak biases into strong constraints on linguistic patterning. In subsequent work, Kirby, Cornish and Smith (2008) confirmed such amplification of weak biases when cultural evolution was implemented by having human learners receive as input what a previous learner had produced as output, thus simulating the cross-generational transmission of a linguistic system. Similarly, a shared sign system can emerge culturally from interactions between artificial (Steels 2003) and human (Fay, Garrod & Roberts 2008) agents. Phylogenetic modeling of existing language patterns relating to word order also points to the cultural evolution of language (Dunn, Greenhill, Levinson & Gray 2011), in line with typological linguistic analyses (Evans & Levinson 2009).

At the theoretical level, cultural evolution has become the centerpiece of a number of accounts of the emergence of linguistic structure (e.g. Christiansen 1994; Christiansen & Chater 2008; Deacon 1997; Kirby & Hurford 2002). For example, Tomasello (2000) suggested that "… the actual grammatical structures of modern languages were humanly created through processes of grammaticalization during particular cultural histories, and through processes of cultural learning, …" (p. 163). An implication of the cultural evolution view is that much of the neural hardware involved in language is not specific to it (though see Christiansen et al. 2011, regarding possible biological adaptations for functional features of language). That is, language has to be acquired largely by mechanisms that are not uniquely dedicated for this purpose. Crucially, though, the cultural evolution perspective does not deny the existence of genetic constraints on language but instead questions the presupposition that these necessarily have to be linguistic in nature.

But how could language have evolved to be as complex as it is without language-specific constraints? In this chapter, I describe computational simulations that provide a possible answer to this question: language has evolved to rely on a multitude of probabilistic information sources for its acquisition, allowing it to be as expressive as possible while still being learnable by domain-general mechanisms (see also Chater & Christiansen 2010; Christiansen & Chater 2008). In the first section, *Language Shaped by Multiple Constraints*, I introduce the evolutionary simulations, from which I derive predictions for specific properties of modern language. Next, in *Multiple-Cue Integration in Language Acquisition and Processing*, I explore predictions from the simulations regarding the role of phonological and distributional information in syntactic acquisition and processing. Finally, I discuss the usefulness of computational modeling in theorizing about language evolution in *Modeling the Cultural Evolution of Language*.

2. Language shaped by multiple constraints

The cultural evolution perspective suggests that the structure of language derives primarily from processes of cultural transmission involving repeated cycles of learning and use. Thus, similar to the proposed cultural recycling of prior cortical maps for recent human innovations such as reading and arithmetic (Dehaene & Cohen 2007), I suggest that language likewise has evolved by "piggybacking" on pre-existing neural substrates, inheriting their structural constraints. These constraints – including socio-pragmatic considerations, the nature of our thought processes, perceptuo-motor factors, as well as cognitive limitations on learning, memory, and processing – have subsequently been amplified and embedded in language through cultural evolution (Christiansen & Chater 2008). In this way, cultural transmission – both vertically (across generations) and horizontally (within generations) – has shaped language to be as learnable and processable as possible given multiple constraints deriving from the human brain. But might this process of multiple-constraint satisfaction have ramifications beyond cultural evolution, affecting how we acquire and use language? Here, I explore the hypothesis that pre-existing neural constraints not only provided important restrictions on cultural evolution but also made available multiple sources of information – or *cues* – that can facilitate both the acquisition and use of language. By "recruiting" such cues, some of which may be partially overlapping and redundant, language could evolve culturally to become more expressive, while still being learnable and processable by mechanisms not dedicated to language.

The proposal put forward here is thus that cultural evolution has shaped language to depend on multiple-cue integration for its acquisition and processing. To provide an initial existence proof of how this might work, I first present results from evolutionary simulations indicating that cultural evolution can lead to the recruitment of cues to facilitate acquisition. A prediction from this simulation is that every language today, as a product of cultural evolution, should incorporate its own unique constellation of probabilistic cues to signal different aspects of linguistic structure. Indeed, prior research has found evidence of such multiple-cue integration across different levels of linguistic representation, from word segmentation (e.g. Mattys, White & Melhorn 2005) to syntactic relations (Monaghan & Christiansen 2008) and beyond (Evans & Levinson 2009).

2.1 Recruitment of cues during cultural evolution of language

Christiansen and Dale (2004) conducted a set of simulations to investigate whether cultural evolution could result in the recruitment of cues to facilitate the learning of more complex linguistic structure. As learners, they employed simple recurrent networks (SRNs, Elman 1990), a type of connectionist model that implements a domain-general learner with sensitivity to complex sequential structure in the input. A crucial feature of this model is that it is self-supervised: it learns from predicting the next element in a sequence and thus learns from comparing its predictions to what actually occurs next. This model has been successfully applied to the modeling of both language processing (e.g. Elman 1993) – including multiple-integration (Christiansen, Allen & Seidenberg 1998; Christiansen, Dale & Reali 2010) – and sequential learning (e.g. Botvinick & Plaut 2004). As a model of human performance, the SRN has been shown to closely mimic the processing of different kinds of recursive linguistic constructions (Christiansen & Chater 1999; Christiansen & MacDonald 2009) as well as the sequential learning of nonadjacent dependencies (Misyak, Christiansen & Tomblin 2010). In addition, the SRN has been applied to the modeling of potential coevolution between language and learners (Batali 1994).

The languages, on which the models were trained, were generated by small context-free grammars, each derived from the grammar skeleton illustrated in Table 1. The curly brackets indicate that the order of the constituents on the right-hand side of a rule can be either as is (head-first) or in the reverse order (head-final). The SRNs were expected to use the distributional information afforded by the order of words in the sentences as a cue to the underlying structure of the language. As additional cues to linguistic structure, the languages could recruit a constituent cue and a lexical cue. The constituent cue was an additional input unit that could mark phrase boundaries by being activated following the constituents from a particular phrase structure rule (e.g. N (PP) #, where "#" indicates the activation of the constituent cue after the NP and optional PP). The lexical cue was another input unit that could be coactivated with any of the 24 words in the vocabulary. Thus, there were three potential sources of information for learning about the structure of a language in the form of distributional, constituent, and lexical cues.

Table 1. The grammar skeleton used by Christiansen and Dale (2004)

S	→	{NP VP}
NP	→	{N (PP)}
NP	→	{N PossP}
VP	→	{V (NP)}
PP	→	{adp NP}
PossP	→	{poss NP}

Cultural evolution was simulated by having five different languages compete against one another, with fitness determined by how easy it was for the SRNs to learn a language. At the beginning of a simulation, five different languages were randomly generated based on the grammar skeleton with a random combination of constituent and lexical cues. Each language was then learned by five different SRNs, with a language's overall fitness being computed as the average across the five networks. The most easily learned language, along with four variations of it, would then form the basis for the next generation of languages, each being learned by five networks. Again, the most easily learned language would be selected as the parent of the next generation, and the process repeated until the same language won for 50 consecutive generations. Language variation was implemented by randomly changing two of the three cues: (1) changing the head-order of a rule, (2) adding or deleting the constituent unit for a rule, or (3) adding or deleting the coactivation of the lexical unit for a word. Ten different simulations were run, each with different initial randomizations.

Of the ten simulations, one never settled but the results of the remaining nine simulations followed a consistent pattern. First, all languages ended up with a highly regular head-ordering, with at least five of the six phrase structure rules being either all head-initial or all head-final. This fits the general tendency for word order patterns in natural languages to be either head-initial or head-final (e.g. Dryer 1992). Second, the constituent cue always separated NPs from other phrases, consistent with evidence from corpus analyses indicating that prosodic cues, such as pauses and pitch changes, are used to delineate phrases in both English and Japanese child-directed speech (Fisher & Tokura 1996). Finally, the lexical cue reliably separated word classes, with six of the runs resulting in the lexical cue separating function words from content words. This is similar to the acoustic differentiation of function and content words observed in English (Cutler 1993). To place these results in context, it is important to note that given the combination of the three different cues in these simulations, there were nearly three-quarters of a million[1] different possible linguistic systems that could have evolved through cultural evolution. Thus, it is not a trivial matter that these simulations culminated in linguistic systems that incorporate properties closely resembling those of natural language.

The simulations by Christiansen and Dale (2004) suggest that linguistic systems can recruit cues to facilitate learning when undergoing cultural evolution. The integration of these cues, in turn, allows language to become more complex while still

1. The number of possible linguistic systems was calculated as follows: there were 6 rules with 2 head-orderings, each with or without the constituent cue, and with each language having a 24-word vocabulary, in which each word could be associated with the lexical cue or not: $6^4 \times 24^2 = 746{,}496$.

being learnable by domain-general mechanisms. Because these cues are probabilistic in nature, and therefore unreliable when considered in isolation, multiple-cue integration has become a necessary component of language acquisition. This gives rise to the following prediction: if natural language has evolved culturally to rely on multiple-cue integration, as indicated by these simulations, then it should be possible to (1) uncover useful cues in modern languages, (2) show computationally that these cues can facilitate language acquisition, and (3) demonstrate that children and adults actually utilize such cues for acquisition and processing. The next section details studies designed to test these predictions from the evolutionary simulations.

3. Multiple-Cue integration in language acquisition and processing

The evolutionary simulations indicated that cultural evolution can lead to the recruitment of cues to facilitate acquisition. In the present section, I review empirical results from studies aimed to test the three predictions derived from the evolutionary simulations, focusing on two of the cues explored in those simulations: lexical information in the form of the phonology of individual words and distributional information in the form of word co-occurrence patterns. First, quantitative results from corpus analyses demonstrate that phonological and distributional cues can be integrated to provide reliable information about lexical categories. Computational modeling then confirms that domain-general learners can successfully utilize such cues. Finally, human experiments establish that children use phonological cues during word learning and that adults cannot help but pay attention to the sound of syntax during sentence processing.

3.1 Quantifying the usefulness of phonological and distributional cues

Do the phonological forms of words contain information relevant for syntax acquisition? The standard assumption of the arbitrariness of the sign (de Saussure 1916) might be taken to suggest otherwise. Indeed, it has been argued that it is a universal characteristic of human language that the relationship between the form of a word and its meaning is arbitrary (Hockett 1960). This assumption is fundamental to most modern grammatical theories on both sides of the Chomskyan divide. For example, Pinker (1999, p. 2) states that "onomatopoeia and sound symbolism certainly exist, but they are asterisks to the far more important principle of the arbitrary sign – or else we would understand the words in every foreign language instinctively, and never need a dictionary for our own!" In a similar vein, Goldberg (2006, p. 217) notes that "… the particular phonological forms that a language chooses to convey particular concepts […] generally are truly arbitrary, except in relative rare cases of phonaesthemes." However, the simulations by Christiansen and Dale (2004) suggest that the sign may not be entirely arbitrary; rather, a systematic relationship should exist between the sound of a word and its lexical category, if the phonological form of a word is to be useful for syntax acquisition as a lexical cue.

In a series of corpus analyses of child-directed speech, Monaghan, Chater, and Christiansen (2005) quantified the potential usefulness of phonological cues to lexical categories. More than five million words were extracted from the CHILDES database (MacWhinney 2000), comprising more than a million utterances spoken in the presence of children. Phonological forms and lexical categories were gleaned from the CELEX database (Baayen, Pipenbrock & Gulikers 1995) and results reported for the 5,000 most frequent words. As potential cues to lexical categories, Monaghan et al. used 16 different phonological properties (listed in Table 2) that have been proposed to be useful for separating nouns from verbs (and function words from content words). Instead of treating each cue in isolation, the 16 cues were combined into a unified

Table 2. The 16 phonological cues used by Monaghan, Chater, and Christiansen (2005)

Phonological cue	Example: *fingers*
Word level	
Length in phonemes	6
Length in syllables	2
Presence of stress	1
Syllable position of stress	1
Syllable level	
Number of consonants in word onset	1
Proportion of phonemes that are consonants	0.66
Proportion of syllables containing reduced vowel	0.5
Reduced first vowel	0
-*ed* inflection	0
Phoneme level	
Proportion of consonants that are coronal	0.25
Initial /ð/	0
Final voicing	1
Proportion of consonants that are nasals	0.25
Position of stressed vowel	1
Position of vowels	1.5
Height of vowels	1

phonological representation for each word. A statistical analysis[2] was then conducted using these representations, resulting in decent classifications of both nouns (58.5%) and verbs (68.3%) – with an indication that phonological cues may be more useful for discovering verbs than nouns. The advantage of phonological cues for verbs was subsequently confirmed by further analyses in Christiansen and Monaghan (2006).

Importantly, though, because the phonological cues are probabilistic in nature, they did not predict the lexical category of a word perfectly. As an additional cue, Monaghan et al. therefore assessed the usefulness of distributional information using a simple, developmentally plausible approach. They selected the 20 most frequent words in the corpus (*are, no, there, this, your, that's, on, in, oh, do, is, and, I, that, what, to, a, it, the, you*) and recorded how often these preceded one of the target words (e.g. *you want*). The rationale was that, even though the child may not know the meaning of the 20 context words, these word forms nonetheless constitute highly frequent acoustic events to which the child is likely to be sensitive. To determine the usefulness of the distributional patterns thus recorded, Monaghan et al. used an information-based measure[3] to assess the strength of the association between the context word and the target word. The distributional information relevant to a specific target word was then represented as a unified representation containing the associative strength scores for each of the 20 context words (e.g. for *cats*, the scores for *are cats, no cats, there cats,* and so on). These distributional cue representations were then submitted to the same statistical analysis as the phonological cues. The results showed a very good classification of nouns (93.7%) but not of verbs (31.1%).

The results from the two analyses suggest that the usefulness of phonological and distributional cues may differ across nouns and verbs. Perhaps integration across the two types of cues may improve classification? Monaghan et al. combined the

2. Each word was represented by a 16-place vector. This means that each word correponds to a point in a 16-dimensional space defined by the 16 phonological cues. Monaghan et al. employed a discriminant analysis to determine whether the nouns and verbs formed separate clusters in this phonological space. Informally, this type of statistical analysis inserts a hyperplane into the 16-dimensional phonological cue space to produce the optimal separation of nouns and verbs into two different categories. Correct classification of nouns and verbs can then be computed given how well the hyperplane separates the two categories from one another.

3. Monaghan et al. used an adapted version of the Dunning (1993) log-likelihood score to estimate the informational value of the distributional cues. Informally, this measure provides an estimation of how surprising it is that the context and target words occur together given how often each occurs on its own. Each word was then represented as a 20-place vector, corresponding to the signed Dunning log-likelihood scores for each of the 20 context words. Classification of nouns and verbs given these distributional representations was then assessed using a discriminant analysis.

phonological cues and the distributional cues into a combined representation,[4] redid their analyses, and obtained reliable classifications of both nouns (67.0%) and verbs (71.4%). When considering correct classifications of nouns and verbs together, they further noted an interesting interaction of phonological/distributional cues with frequency, as shown in Figure 1. Distributional cues (gray bars) appear to work very well for high-frequency words but less so for low-frequency words. This is likely because high-frequency words occur in more contexts and this provides for more accurate distributional information about their lexical categories. Phonological cues (white bars), on the other hand, seem to work better for low-frequency words than for high-frequency words. This may be explained by the tendency for high-frequency words to be shortened, perhaps leading to the omission of important phonological cues to their lexical category. In contrast, low-frequency words are not subjected to the same shortening pressures, allowing the cues to remain in place. Crucially, though, when the two types of cues are integrated, good classification can be found across all frequency bins, as illustrated in Figure 1 (black bars).

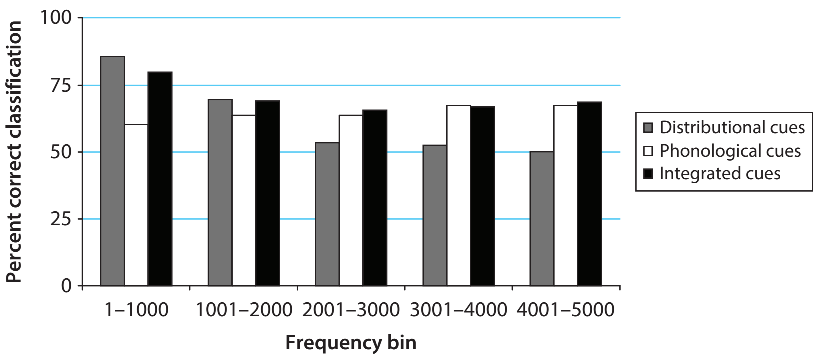

Figure 1. The percentage of nouns and verbs correctly classified as such across different frequency bins for distributional (gray bars) and phonological cues (white bars) treated separately, and when both cues are integrated with one another (black bars)

The results presented so far apply only to English. If language, in general, has evolved to rely on multiple-cue integration, then it should be possible to find evidence of similar kinds of cue information in other languages as well. However, many of phonological cues used by Monaghan et al. (2005) were specific to English (Table 1)

4. This simply involved combining the 16-place phonological cue vector with the 20-place distributional cue vector, resulting in a 36-place multiple-cue vector representation for each word. A discriminant analysis was conducted on the 36-dimensional cue space defined by these word representations.

and thus may not work for other languages. Monaghan, Christiansen, and Chater (2007) therefore generated a set of 53 cross-linguistic phonological cues, including gross-level word cues such as length, consonant cues relating to manner and place of articulation of phonemes in different parts of the words, and vowel cues relating to tongue height and position as well as whether the vowel was reduced. They then conducted analyses of child-directed speech in English, French, and Japanese. Using the new cues, they replicated the results of the previous study in terms of correct noun/verb classification (16 cues: 63.4% vs. 53 cues: 67.5). Noun/verb classification using phonological cues was also very good for both French (82%) and Japanese (82%). Classification performance was further improved across all three languages (English: 94%; French: 91.4%; Japanese: 93.4%) when the phonological cues were integrated with distributional cues (computed as before).

Together, the results of these corpus analyses show that, across representatives of three different language genera – Germanic (English), Romance (French), and Japanese – child-directed speech contains useful cues for distinguishing between nouns and verbs (see also Kelly 1992). The results are thus consistent with the prediction that, as a result of the cultural evolution of language, words contain within them the sound of syntax: nouns and verbs differ in terms of their phonology.[5] Importantly, the specific cues differed considerably across languages, suggesting that each language has recruited its own unique set of cues to facilitate acquisition through multiple-cue integration. However, these analyses only demonstrate that there are probabilistic cues available for learning about aspects of syntax. Next, we shall see that a domain-general sequential learner, the SRN, can take advantage of both phonological and distributional cues to learn about syntax.

3.2 Multiple-cue integration by a sequential learner

A potential concern regarding multiple-cue integration is that there are many kinds of information that could potentially inform language acquisition. As noted by Pinker (1984, p. 49),

> … in most distributional learning procedures there are vast numbers of properties that a learner could record, and since the child is looking for correlations among these properties, he or she faces a combinatorial explosion of possibilities. … Adding semantic and inflectional information to the space of possibilities only makes the explosion more explosive.

5. That the phonological forms of words carry information about their syntactic use as nouns or verbs does not necessarily require the postulation of universal lexical categories. Instead, phonological and distributional cues provide probabilistic information about how words can be used in sentential contexts and this is what is assessed by the corpus analyses reported in this chapter.

Pinker expresses a common intuition about the use of multiple, partially correlated sources of information by a domain-general learning device: that combining different kinds of partially reliable information can only result in unreliable outcomes. However, research in formal learning theory has shown that this intuition is incorrect. Mathematical analyses of neural network learning using the Vapnik-Chervonenkis Dimension[6] have shown that multiple-cue integration with correlated information sources will not lead to a combinatorial explosion but instead to improved learning (Abu-Mostafa 1993). This holds even when one or more of the cues are either uncorrelated or otherwise uninformative with respect to the acquisition task, in which case they have no negative effect on learning (see Allen & Christiansen 1996, for neural network applications, including to the SRN). Thus, mathematically speaking, Pinker's intuitive concern about combinatorial explosion is unfounded.

Although the issue of combinatorial explosion is not a problem in principle, it may nonetheless pose a considerable obstacle in practical terms. Christiansen and Dale (2001) sought to address this issue head-on by training SRNs to do multiple-cue integration, given a corpus of artificially generated child-directed speech. The corpus incorporated declarative, imperative, and interrogative sentences with subject-noun/verb agreement and variations in verb argument structure. In one simulation, the networks were provided with three partially reliable cues to syntactic structure (word length, lexical stress, and pitch change) and three cues not related to syntax (presence of word-initial vowels, word-final voicing, and relative speaker pitch). The results of the simulations indicated that the SRNs were able to ignore the three unrelated cues while taking full advantage of informative ones, as indicated by the mathematical analyses.

The question remains, though, whether Christiansen and Dale's SRN model can scale up to deal with the kind of cues found in the corpus analyses described previously. To answer this question, Reali, Christiansen, and Monaghan (2003) trained SRNs on a full-blown corpus of natural speech directed at children between the ages of 1 year and 1 month to 1 year and 9 months (Bernstein-Ratner 1984). Each word in the input was encoded in terms of the 16 phonological cues used in the Monaghan et al. (2005) corpus analyses (and shown in Table 2). Given a word represented in terms of these phonological cues, the task of the networks was to predict the next lexical category in the utterance. Thus, the network would receive both phonological cues (in terms of the 16-cue representations for each word) and distributional cues (in terms of the co-occurrence of words in the corpus). To assess the usefulness of

6. The Vapnik-Chervonenkis (VC) dimension establishes an upper bound for the number of examples needed by a learning process that starts with a set of hypotheses about the task solution. A hint may lead to a reduction in the VC dimension by weeding out bad hypotheses and reduce the number of examples needed to learn the solution.

the phonological cues relative to the distributional ones, a second group of networks was also trained. For these networks, the phonological-cue representation for a given word was randomly reassigned to a different word to break the correlations between phonology and lexical category. The results showed that the SRNs trained to integrate phonological and distributional cues performed significantly better than the networks provided only with distributional cues. Further analyses of the networks' internal states indicated that the phonological cues were particularly useful for processing novel words, allowing the network to place itself in a "noun state" when processing novel nouns and in a "verb state" when encountering new verbs.

The results of the SRN simulations indicate that a domain-general sequential learner can learn aspects of syntactic structure via multiple-cue integration, as predicted by the evolutionary simulations. Despite intuitions to the contrary, a combinatorial explosion does not occur. Rather, the right cues are recruited to facilitate acquisition because the language has evolved to be learnable by way of those very cues. For example, phonological cues promote better learning and better generalization to new words. To be able to take advantage of these cues, children become attuned to the relevant cues in their native language during the first years of life, as we shall see next.

3.3 Phonological cues in acquisition and processing

The corpus analyses indicated that there are useful phonological cues for language acquisition and the SRN simulations demonstrated that a sequential learner can take advantage of them – but are children sensitive to phonological cues when exposed to new words? Storkel (2001, 2003) has shown that preschoolers find it easier to learn novel words when these consist of phonotactically common sound sequences. However, these studies did not address the question of whether children may use phonological cues to learn about the syntactic role of words. Fitneva, Christiansen, and Monaghan (2009) therefore conducted a word learning study to investigate whether children implicitly use phonological information when guessing about the referents of novel words. To create novel words that were either noun-like or verb-like in their phonology, Fitneva et al. used a measure of phonological typicality, originally proposed by Farmer, Christiansen, and Monaghan (2006). Phonological typicality measures how typical a word's phonology is relative to other words in its lexical category, and reliably reflects the phonological coherence of nouns and verbs (Monaghan, Christiansen, Farmer & Fitneva 2010). Thus, noun-like nouns are typical in terms of their phonology of the category of nouns, and likewise verb-like verbs are phonologically typical of other verbs. When asking English monolingual second-graders to guess whether a novel word referred to a picture of an object or a picture of an action, the children used the phonological typicality of the nonword in making their choices. Interestingly, as predicted by the corpus analyses (Christiansen & Monaghan 2006), verbs benefited more from phonological cues than nouns.

It may be objected that second-graders – as language learners – are too “old” to serve as a suitable population with which to investigate the usefulness of phonological cues, especially if such cues are to be used to inform early syntactic acquisition. To address this objection, Fitneva et al. conducted a second study with another group of second-graders, who were enrolled in a French immersion program. The stimuli were the same as in the experiment with the monolingual children. Crucially, though, whereas half the nonwords were verb-like and the other half noun-like with respect to English phonology, all the nonwords were noun-like according to French phonology. Two groups of the English-French bilingual children were tested, with the only difference being in the language used for the instructions. When given English instructions, the bilingual children behaved exactly like the monolingual English children, showing an effect of English phonological typicality. However, when the instructions were provided in French, the patterns of results changed, in line with French phonology. Hence, not only did the children seem to use phonological cues to make guesses about whether a novel word is a noun or a verb but they were also able to do so after relatively little experience with the relevant phonology (less than two years of exposure in a formal educational setting for the children in the French immersion program).

The results of the word learning study suggest that phonological cues may come into play early in syntax acquisition. Farmer et al. (2006) explored whether multiple-cue integration involving phonological cues extends into adulthood. Using the measure of phonological typicality, they demonstrated that the processing of words presented in isolation is affected by how typical their phonology is relative to their lexical category: noun-like nouns are read aloud faster, as are verb-like verbs. Similarly, Monaghan et al. (2010) showed that people are faster to make lexical decisions about whether a presented item is a real word or not, if that word is phonologically typical of its lexical category. Farmer et al. further showed that the phonological typicality of a word could even affect how easy it is to process in a sentence context. Indeed, for noun/verb homonyms (e.g. *hunts* as in *the bear hunts were terrible…* versus *the bear hunts for food…*), if the continuation of the sentence is incongruent with the phonological typicality of the homonym, then people both experience online processing difficulties and have problems understanding the meaning of the sentence.

Together, the results of the human experimental studies confirm the prediction from the evolutionary simulations, indicating that the use of phonological cues during acquisition is so important that it becomes a crucial part of the developing language processing system. The phonological properties of words facilitate lexical acquisition through multiple-integration and become an integral part of lexical representations. As consequence, adult language users cannot help but pay attention to phonological cues to syntactic structure when processing language.

4. Modeling the cultural evolution of language

In this chapter, I have proposed that language has evolved by way of cultural evolution to exploit multiple cues so as to be maximally expressive while still being learnable by domain-general mechanisms. Evolutionary simulations were discussed, indicating how language may recruit cues to facilitate learning. One prediction from this perspective on the cultural evolution of language is that each of the world's languages should have its own constellation of cues. Cross-linguistic corpus analyses have confirmed this prediction with regard to phonological cues, showing that the relationship between a word's sound and how it is used is not arbitrary. Computational modeling substantiated a second prediction from the evolutionary simulations: that domain-general sequential learners should be able take advantage of phonological cues in the context of multiple-cue integration. Finally, evidence from human experimentation showed that children use phonological cues during word learning, as do adults during sentence processing, corroborating a third prediction from the evolutionary simulations suggesting that the use of phonological cues becomes a crucial part of the emerging language processing system. Of course, phonological cues are not the only useful sources of information for learning about aspects of syntax; rather, they are integrated with other sources of information during language acquisition, including distributional (e.g. Redington, Chater & Finch 1998), prosodic (e.g. Fisher & Tokura 1996), semantic (e.g. Bowerman 1973), and pragmatic (e.g. Tomasello 2003) cues (see Monaghan & Christiansen 2008; Morgan & Demuth 1996, for reviews). Hence, language has evolved to rely on multiple-cue integration in both acquisition and processing, making it integral to the computational architecture of our language system.

More generally, the evolutionary simulations illustrate how computational modeling may inform theories of language evolution by providing the means for evaluating current theories, exploring new theoretical constructs, and/or offering existence proofs that specific hypotheses could work (Christiansen & Kirby 2003). First, the simulations constitute an explicit *evaluation* of the degree to which subtle learning biases (inherent in the SRNs; Christiansen & Chater 1999) can drive the cultural evolution of linguistic structure (here, toward word order regularities). Second, the simulations *explore* how multiple cues may interact to facilitate the cultural evolution of language. Third, the simulation results provide an *existence proof* that a culturally evolving linguistic system can recruit cues to make itself easier to learn and process by domain-general learners. Moreover, the evolutionary simulations also made it possible to derive empirical predictions about extant language, and these were subsequently confirmed by empirical work involving corpus analyses, computational modeling, and human experimentation.

On a more theoretical level, construing language evolution as primarily involving cultural evolution may offer a possible counterpoint to Lewontin's (1998) scathing critique of evolutionary approaches to cognition, and to language evolution in particular: "Reconstructions of the evolutionary history and the causal mechanisms of the acquisition of linguistic competence […] are nothing more than a mixture of pure speculation and inventive stories" (p. 111). Focusing on evolutionary psychology (e.g. Pinker & Bloom 1990), he doubts that heritable variation in linguistic abilities among individuals in the hominid lineage would ever be associated with having more offspring. His main concern is that it is impossible to test the hypotheses put forward to explain biological evolution of language because of our limited knowledge about hominid evolution in general. However, the kind of cultural evolution perspective espoused here does not seem to suffer from this problem because, as illustrated in this chapter, we can test specific evolutionary scenarios through computational simulations from which we can derive testable predictions that can be substantiated through empirical research. In the specific case discussed in this chapter, the results suggest that cultural evolution has shaped language to depend on multiple-cue integration.

Acknowledgments

Sections 2 and 3 in this chapter are adapted with permission from the section *The Sound of Syntax* in the chapter *Language Has Evolved to Depend on Multiple-Cue Integration* in the volume *The Evolutionary Emergence of Language*, edited by Rudolf Botha and Martin B.H Everaert for Oxford University Press. Many thanks to Christina Behme for helpful comments on these two sections in their original incarnation.

References

Abu-Mostafa, Y.S. (1993). Hints and the VC Dimension. *Neural Computation, 5*, 278–288.

Allen, J., & Christiansen, M.H. (1996). Integrating multiple cues in word segmentation: A connectionist model using hints. In *Proceedings of the* 18th *Annual Cognitive Science Society Conference* (pp. 370–375). Mahwah, NJ: Lawrence Erlbaum Associates.

Baayen, R.H., Pipenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks & P. Maes (Eds.), *Artificial life IV* (pp. 160–171). Cambridge, MA: The MIT Press.

Bernstein-Ratner, N. (1984). Patterns of vowel modification in motherese. *Journal of Child Language, 11*, 557–578.

Botvinick, M., & Plaut, D.C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review, 111*, 395–429.

Bowerman, M. (1973). Structural relationships in children's utterances: Semantic or syntactic? In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 197–213). New York, NY: Academic Press.

Brighton, H., Smith, K., & Kirby, S. (2005). Language as an evolutionary system. *Physics of Life Reviews, 2*, 177–226.

Chater, N., & Christiansen, M.H. (2010). Language acquisition meets language evolution. *Cognitive Science, 34*, 1131–1157.

Chater, N., Reali, F., & Christiansen, M.H. (2009). Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences, 106*, 1015–1020.

Christiansen, M.H. (1994). *Infinite languages, finite minds: Connectionism, learning and linguistic structure*. Ph.D. thesis, University of Edinburgh.

Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes, 13*, 221–268.

Christiansen, M.H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science, 23*, 157–205.

Christiansen, M.H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*, 487–558.

Christiansen, M.H., & Dale, R. (2001). Integrating distributional, prosodic and phonological information in a connectionist model of language acquisition. In *Proceedings of the* 23rd *Annual Conference of the Cognitive Science Society* (pp. 220–225). Mahwah, NJ: Lawrence Erlbaum Associates.

Christiansen, M.H., & Dale, R. (2004). The role of learning and development in the evolution of language. A connectionist perspective. In D. Kimbrough Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach* (pp. 90–109). Cambridge, MA: The MIT Press.

Christiansen, M.H., Dale, R., & Reali, F. (2010). Connectionist explorations of multiple-cue integration in syntax acquisition. In S. P. Johnson (Ed.), *Neoconstructivism: The new science of cognitive development* (pp. 87–108). New York, NY: Oxford University Press.

Christiansen, M.H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences, 7*, 300–307.

Christiansen, M.H., & MacDonald, M.C. (2009). A usage-based approach to recursion in sentence processing. *Language Learning, 59*, 126–161.

Christiansen, M.H., & Monaghan, P. (2006). Discovering verbs through multiple-cue integration. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets words: How children learn verbs* (pp. 88–107). New York, NY: Oxford University Press.

Christiansen, M.H., Reali, F., & Chater, N. (2011). Biological adaptations for functional features of language in the face of cultural evolution. *Human Biology*, *83*, 247–259.

Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research, 22*, 109–131.

de Saussure, F. (1916). *Course in general linguistics*. New York, NY: McGraw-Hill.

Deacon, T. (1997). *The symbolic species: The coevolution of language and the brain*. New York, NY: Norton.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*, 384–398.

Dryer, M. (1992). The Greenbergian word order correlations. *Language, 68*, 81–138.

Dunn, M., Greenhill, S.J., Levinson, S.C., & Gray, R.D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, *473*, 79–82.

Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics, 19*, 61–74.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48*, 71–99.

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*, 429–492.

Farmer, T.A., Christiansen, M.H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences, 103*, 12203–12208.

Fay, N., Garrod, S., & Roberts, L. (2008). The fitness and functionality of culturally evolved communication systems. *Philosophical Transactions of the Royal Society of London (B), 363*, 3553–3561.

Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development, 67*, 3192–3218.

Fitneva, S.A., Christiansen, M.H., & Monaghan, P. (2009). From sound to syntax: Phonological constraints on children's lexical categorization of new words. *Journal of Child Language, 36*, 967–997.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. New York, NY: Oxford University Press.

Hockett, C.F. (1960). The origin of speech. *Scientific American, 203*, 89–96.

Kelly, M.H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349–364.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, 105*, 10681–10685.

Kirby, S., Dowman, M., & Griffiths, T. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences, 104*, 5241–5245.

Kirby, S., & Hurford, J. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the evolution of language* (pp. 121–147). London: Springer.

Lewontin, R.C. (1998). The evolution of cognition: Questions we will never answer. In D.S. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science, Volume 4: Methods, models, and conceptual issues* (pp. 107–131). Cambridge, MA: The MIT Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd. ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Mattys, S.L., White, L., & Melhorn, J.F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General, 134*, 477–500.

Misyak, J.B., Christiansen, M.H., & Tomblin, J.B. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science, 2*, 138–153.

Monaghan, P., Chater, N., & Christiansen, M.H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition, 96*, 143–182.

Monaghan, P., & Christiansen, M.H. (2008). Integration of multiple probabilistic cues in syntax acquisition. In H. Behrens (Ed.), *Trends in corpus research: Finding structure in data* (pp. 139–163). Amsterdam: John Benjamins.

Monaghan, P., Christiansen, M.H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology, 55*, 259–305.

Monaghan, P., Christiansen, M.H., Farmer, T.A., & Fitneva, S.A. (2010). Measures of phonological typicality: Robust coherence and psychological validity. *The Mental Lexicon, 5*, 281–299.

Morgan, J.L., & Demuth, K. (Eds.). (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Nowak, M.A., Komarova, N.L., & Nyogi, P. (2002). Computational and evolutionary aspects of language. *Nature, 417*, 611–617.

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

Pinker, S. (1999). *Words and rules*. New York, NY: Basic Books.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–727.

Reali, F., Christiansen, M.H., & Monaghan, P. (2003). Phonological and distributional cues in syntax acquisition: Scaling up the connectionist approach to multiple-cue integration. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 970–975). Mahwah, NJ: Lawrence Erlbaum Associates.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science, 22,* 425–469.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, *7*, 308–312.

Storkel, H.L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44*, 1321–1337.

Storkel, H.L. (2003). Learning new words II: Phonotactic probability in verb learning. *Journal of Speech, Language, and Hearing Research, 46*, 1312–1323.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4, 156–163.

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