The phonological mind

Iris Berent
Department of Psychology, Northeastern University, 125 Nightingale Hall, 360 Huntington Ave, Boston MA 02115, USA

Humans weave phonological patterns instinctively. We form phonological patterns at birth, we spontaneously generate them de novo, and we impose phonological design on both our linguistic communication and cultural technologies—reading and writing. Why are humans compelled to generate phonological patterns? Why are phonological patterns intimately grounded in their sensorimotor channels (speech or gesture) while remaining partly amodal and fully productive? And why does phonology shape natural communication and cultural inventions alike? Here, I suggest these properties emanate from the architecture of the phonological mind, an algebraic system of core knowledge. I evaluate this hypothesis in light of linguistic evidence, behavioral studies, and comparative animal research that gauges the design of the phonological mind and its productivity.

Some puzzles
All languages construct words (meaningful symbols) from meaningless elements. English speakers contrast gods and dogs, they write blogs not lblogs, and they rhyme them with frogs. Such patterns are not merely reflexes of speech, because similar meaningless structures are found in sign languages [1]. And in both modalities, these patterns generalize to new forms [2,3]. The tacit knowledge of humans concerning the patterning of meaningless linguistic elements is called phonology (see Glossary).

Why do humans weave phonological patterns? Viewed broadly, vocal patterns of meaningless elements are common in nature [4]. But while animal communication is typically attributed to specialized cognitive systems [5], human phonological patterns are viewed as products of domain-general sensorimotor pressures (e.g., blog is easier to perceive and articulate than lblog) and associative mechanisms of statistical learning [6,7]. Even those who endorse the specialization of the language faculty tend to rest their case on syntax [8,9]; phonological specialization is deemed unlikely [10].

But this conclusion is premature (Figure 1) — it fails to explain why distinct phonological systems (signed and spoken) converge on their design [11]; why this design relies on algebraic mechanisms [12,13] (rather than merely analog sensorimotor pressures); why phonology emerges early [14] and spontaneously [15,16]; and why it forms the basis for reading [17]. Here, I suggest that these properties emanate from the architecture of the phonological mind; — the possibility that phonology is an algebraic system of core knowledge. I spell out these two hypotheses below and

Glossary

Algebraic optimization: the capacity to optimize functional pressures (e.g., phonetic restrictions) by relying on algebraic rules (e.g., algebraic phonological rules).

Algebraic rules: mental operations that can extend regularities across the board, to any member of a class, actual or potential. These generalizations are supported by various representational capacities, including the capacity (i) to form equivalence classes; (ii) to operate on entire classes using variables; and (iii) to distinguish types (noun) from individual tokens (dog) [28].

Coarticulation: the overlap (partial or full) in the production of distinct sounds (e.g., consonants and vowels) during speech production.

Core knowledge: knowledge of innate, universal principles that determine an individual’s understanding of the world in early development and shapes the acquisition of mature knowledge systems later in development.

Equivalence class: a class of elements whose members (actual or potential) are all treated alike with respect to a given generalization. For example, the CV syllable rule (a syllable is comprised of ‘any consonant’ followed by ‘any vowel’) treats ‘all consonants’ alike; it applies to either familiar English consonants (e.g., b) or novel ones that are nonnative to English (e.g., ch in Chanukah).

Homology: a common trait that emerges in distinct species through descent from a common ancestor.

Homoplasy: a common trait that independently emerged in distinct species (i.e., it is not a homology).

Metrical phonology: humans’ knowledge concerning the relative prominence of phonological units. Such knowledge, for example, allows humans to identify the prominence contrast between the two syllables in baby (where the initial syllable has greater prominence) and begin (with a prominent final syllable).

Morphemes: linguistic units that convey the meaning or function of a word. For example, the word liked comprises two morphemes, the base like and the ending ed, indicating a past tense.

Phonetics: the system that implements algebraic phonological patterns as concrete, analog sensorimotor programs, either acoustic and oral (in spoken languages) or visual and gestural (in sign languages).

Phonological patterns: a pattern of meaningless linguistic elements.

Phonology: humans’ knowledge concerning the patterning of meaningless linguistic elements in their language, either signed or spoken.

Productivity: the capacity to extend linguistic regularities to novel instances.

Onset: the consonant(s) that occur at the beginning of a syllable (e.g., br/ in block).

Ranking (of constraints): modern linguistic theory [38] asserts that all languages share the same universal set of grammatical constraints, but differ with respect to their relative weight (ranking). For example, the fact that syllables such as /ba/ are allowed in Russian, but not English, is captured by the ranking of two constraints. One constraint bans sonority falls (for simplicity, No-ba), whereas another constraint favors faithfulness to such inputs (i.e., for simplicity, Faith-ba). English ranks the ban on /ba/ above the faithfulness to the input and, consequently, such syllables are untested; Russian exhibits the opposite ranking, so /ba/-type syllables are allowed in this language. This example illustrates how a theory of universal grammar can potentially capture both the similarity across languages and their variability.

Sonority: an abstract phonological property of segments that correlates with their loudness. Louder segments (e.g., vowels) are generally more sonorous than softer ones (e.g., stop consonants, such as b,p). Note, however, that sonority is defined by the phonological structure of a segment (i.e., its feature composition), rather than by its loudness per se. Accordingly, vowels remain more sonorous than stops even if they are both presented in print.

Stem: a word base, used in the formation of complex words. For example, dogs is a complex word, formed by adding the plural morpheme s to the stem dog.

Syllable: a meaningless phonological unit that minimally includes a sonority peak (typically, a vowel, e.g., bee) and optionally, lower-sonority margins (typically, consonants, e.g., bar).

Syntax: humans’ knowledge regarding sentence structure.

References
[1] see front matter
1364-6613/3 – © 2013 Elsevier Ltd. All rights reserved. http://dx.doi.org/10.1016/j.tics.2013.05.004

Corresponding author: Berent, I. (i.berent@neu.edu).
Keywords: phonology; core knowledge; algebraic rules; language universals; universal grammar; animal communication; language evolution; reading.

Trends in Cognitive Sciences, July 2013, Vol. 17, No. 7 319
illustrate some of the relevant evidence. A detailed discussion of this thesis can be found elsewhere [18].

Algebraic phonology: a tale of two masters

Phonological patterns carry a double duty. To ensure that words are transmitted efficiently, phonological patterns must be grounded in the phonetic channel (e.g., the auditory and motor systems). Modern phonological theory has repeatedly shown that phonological patterns make phonetic sense [19–21]. The Consonant-Vowel (CV) syllable (e.g., ba) for instance (a structure prevalent across languages) optimizes coarticulation and enhances the perception of acoustic cues [22,23]. However, phonological systems do not answer to their phonetic master directly [24]. This is because the analog sensorimotor system fails to satisfy a second constraint on phonological design: the need to express large lexicons and keep up with the vast productivity of language [25]—the inventions of faxes, blogs, and tweets. Doing so requires that words are formed by conjoining discrete, meaningless elements according to algebraic rules [2]. Consequently, sound-pattern restrictions, such as the preference for CV syllables, are expressed not phonetically (e.g., as a demand for ‘easy production’), but rather as algebraic rules that form syllables by combining broad categories of ‘consonants’ and ‘vowels’. [26]. Phonological patterns thus attain phonetic goals by relying on algebraic means, a property I call ‘algebraic optimization’.

The autonomy of the phonological system from the phonetic interface is supported by linguistic analysis, showing that phonological systems occasionally betray functional goals that conflict with systematic rules [24,27]. Further evidence that phonology employs algebraic rules is presented by experimental findings.

Algebraic rules operate on variables that stand for entire equivalence classes [28,29]. For example, the \#AAB rule applies to two classes (A and B) of ‘any consonant’, and bans identical consonants from occurring at the left edge (e.g., sisum is illicit). Because this class is open ended, the rule applies to any class member, familiar or novel. Identity restrictions are common in phonology [30], and the \#AAB rule is prevalent in Semitic languages; these languages ban identical consonants from the left edge of stems (e.g., sisum) but allow them at their right ends (e.g., simum) [31]. Remarkably, humans generalize this rule not only to novel stems [32–34], but even to ones with nonnative consonants that comprise features that are unattested in the language [12,13] (Box 1). Computational modeling [13,29,35] suggests that such generalizations cannot be
Box 1. The role of algebraic rules in phonology

Many phonological restrictions specifically target identical elements. The example in Figure IA illustrates one such restriction in Hebrew. Hebrew allows identical consonants to occur at the right edge of stems (ABB, e.g., ginun) but disallows them at the left edge (AAB, e.g., gigun). Identity restrictions, by definition, apply to equivalence classes (e.g., ‘any consonant’). To represent such restrictions, it is thus necessary to encode algebraic variables that bind the occurrence of elements across categories (Figure IB): the element selected for the first A category (e.g., the consonant b) must be identical to the one selected for the second A category (i.e., another b). Weaker computational mechanisms that lack operations over variables (certain connectionist networks and Maximum-entropy models) can learn to distinguish AAB from ABB forms, and even extend generalizations to novel items [13,29]. However, these mechanisms fail to extend the identity rule to novel items that are dissimilar to the familiar items experienced by learners (i.e., ones with novel segments and novel features) [13,29,35].

Experimental results indicate that such rules form part of natural phonological systems. The critical evidence is presented by the scope of phonological generalizations (Figure IC). If humans do, in fact, encode an algebraic rule of the form *AAB, then they should be able to generalize their knowledge across the board, not only to the native consonants of their language (e.g., the Hebrew phoneme b) [32–34], but also to nonnative consonants (e.g., th, a phoneme that does not occur in Hebrew, and its place of articulation feature, the wide value of the tongue-tip-constriction-area feature, is likewise non-native). For example, Hebrew speakers should consider thithuk (an illicit AAB form with the nonnative reduplicant th) as worse formed than kithuth (a licit ABB form). In accord with this prediction, Hebrew speakers consider novel AAB stems less acceptable than ABB stems, and they systematically extend this generalization to reduplicant phonemes that are nonnative to the language (Figure ID) [12]. Likewise, because ill-formed stems are less word-like, AAB nonwords are more readily classified as nonwords (in lexical decision) even when they include nonnative phonemes (e.g., thithuk [Figure ID]).

attained by various nonalgebraic mechanisms (i.e., those that lack the capacity to operate on variables that stand for equivalence classes). As such, these results suggest that the phonological mind has the principled capacity to extend generalization across the board, the hallmark of linguistic productivity viewed typically as reserved to syntax [36].

Abstract algebraic rules, such as the *AAB rule, are nonetheless functionally motivated, as the distinction between AAB and ABB forms exacts specific neural costs, additional to those required to encode identity [37]. By eliminating the AAB/ABB contrast, Semitic languages avoid the costly neural operation of binding identity to word edges. Functional costs, however, are addressed indirectly, by relying on algebraic phonological rules that are phonetically motivated. Accordingly, phonology exhibits algebraic optimization.

Phonology is a system of core knowledge

Beyond their capacity to encode algebraic rules, phonological systems further share specific rules [38,39] whose precursors are evident in preverbal infants [14] and nonhuman animals [40–43]. Universality and early onset are hallmarks of core knowledge [44,45]—innate principles that shape both early knowledge and mature cultural inventions. Just as our rudimentary understanding of number [45,46], object [47], agency [48], and living things [49] shapes later theories of mathematics, physics,
morality, and biology, so does phonology form the basis of reading [17,50]. Such similarities suggest that phonology is a system of core knowledge. I review some of the evidence below.

**Shared design**

Linguistic theory suggests that distinct phonological systems exhibit shared representational primitives and constraints. Concerning primitives, all spoken languages encode phonological features; features give rise to segment classes, most notably, consonants and vowels, which combine to form syllables. Experimental findings have implicated these primitives in early development [51]. For example, consonants and vowels have been shown to carry different roles in early language processing (consonants mediate lexical access, whereas vowels convey grammatical information) [52]. Consonants and vowels further segregate neurally [53] and doubly dissociate in aphasia [54].

This inventory of common phonological primitives is combined according to grammatical phonological constraints that are shared across languages [38]. Because these constraints are conflicting and violable, and their relative priorities vary, languages differ on the range of structures they allow (e.g., lba is allowed in Russian, but not in English). Nonetheless, grammatical constraints are putatively universal: they are active in the grammars of all speakers, irrespective of whether the relevant structures are present or absent in their language.

The universality of linguistic knowledge, in general, and phonological constraints, in particular, has been the subject of much recent controversy [55]. Moreover, universality is not itself evidence for specialized core knowledge. Some phonological preferences could be broadly shared because they are guided by generic computational and perceptual principles [56]. The metrical preferences governing the grouping of syllables into prominence patterns are a case in point. Whereas the strong–weak metrical pattern (e.g., **bab**y) is favored over the weak–strong pattern (e.g., **begi**n) [57], and its extraction is preferentially tuned to specific acoustic cues (pitch and intensity) in both adults and infants [58–61], similar preferences occur in the grouping of nonlinguistic musical tones [58,62] and visual shapes [63]. Accordingly, these preferences are likely to reflect principles that are neither algebraic nor specific to phonology.

A stronger case of core phonological principles is presented by the restrictions on syllable structure. Consider, for example, the restrictions on onset clusters (e.g., **bl**ock). Across languages, syllables such as **bla** are preferred (e.g., more frequent) relative to **bna**, which in turn is preferred to **bda**; least preferred is **lba** [64]. These observations are captured by putatively universal constraints known as sonority restrictions [65,66] (Box 2). Unlike metrical prominence, sonority restrictions have no obvious nonlinguistic analog. Although they are functionally motivated (i.e., they optimize speech production and perception) [22,23], sonority restrictions can be formally captured by principles that are algebraic and abstract [65,66]. If such restrictions are effectively active in all grammars, then all speakers should converge on the same preferences (e.g., **bla**–**bna**–**bda**–**lba**) even if these onsets are unattested in their language.

Experimental results from English [64,67–70], Spanish [71], French [72], and Korean [73] support this proposal. The evidence comes from perceptual illusions. Syllables with ill-formed onsets (e.g., **lba**) tend to be systematically misidentified (e.g., as **leb**a)—the worse formed the syllable, the more likely the misidentification. Thus, misidentification is most likely in **lba** followed by **bda**, and is least likely in **bna**. Crucially, the sensitivity to syllable structure occurs even when such onsets are unattested in participants’ languages, and it is evident in adults [64,67–70,73] and young children [74].

Misidentification, of course, could also occur for numerous reasons unrelated to phonology. But auxiliary analyses demonstrate that misidentification is not due to artifacts of the auditory stimuli (Russian speakers, whose language allows those onsets, can identify the same stimuli accurately [64,67]) or to an inability to encode their phonetic form (English speakers misidentify even printed onsets [67,69]). Likewise, the superior identification of better-formed onsets (e.g., **bna**) does not result from analogy to attested onsets (e.g., **bna** resembles **bla**), because similar results obtain in Korean [73], a language that arguably lacks onset clusters altogether. Computational simulations using state-of-the-art inductive learners demonstrate that, absent innate sonority biases, these restrictions are unlearnable [75]. Together, these results suggest that these illusions do not result from sensorimotor pressures or inductive learning. By elimination, then, these findings implicate a grammatical phonological process that recodes ill-formed syllables as better-formed ones. Crucially, these syllable-structure restrictions are broadly shared across languages.

**Spontaneous regenesis: evidence from signed phonologies**

The possibility that phonological systems comprise shared principles does not mean that their design is experience independent. Indeed, absent experience with speech, Deaf individuals acquire sign languages whose structure is different from spoken languages [15]. Nonetheless, all languages, signed and spoken, exhibit a phonological system inasmuch as: (i) they all construct (meaningful) words from patterns of meaningless elements [1,76]; and (ii) the resulting systems share representational primitives and constraints. Like spoken languages, signed languages encode phonological features [77,78], they represent syllables (a representational primitive) and constrain their profile: syllables require a single peak of phonetic energy (either a vowel (a peak of acoustic energy in speech [65]) or a path movement (a peak of visual energy in sign [79]); Figure 2), and this restriction is enforced by signers and nonsigners alike [80].

Crucially, sign languages complete with phonological rules have been shown to emerge spontaneously in Deaf communities that lack experience with sign languages; the Al-Sayyid Bedouin Sign Language (ABSL) is one case in point [15]. However, phonology does not emerge instantaneously, as the first generation of ABSL signers lacked a phonological system. But within three generations, signers produced arbitrary signs that respect putatively universal phonological constraints, including sonority restrictions.
The absence of phonology in the early generations shows that phonological organization is neither physically necessary nor experience independent. Nonetheless, phonological patterns emerge spontaneously in human communities. Spontaneous regenesis is a property shared by human phonological systems and birdsong. In both cases, isolated individuals exhibit an abnormal ‘phonological’ system [81]. Given the necessary triggers, however, both birds [81] and humans [15] eventually converge on patterns that are typical of their conspecifics, but absent in their own experience.

**Ontogeny, phylogeny, and technology**

Whereas all spoken languages exhibit aural phonological patterns, vocal patterns of communication are by no means unique to humans. One thus wonders whether phonology
has homologies or homoplasies in nonhuman communication. To allow us to detect such similarities and control for superficial physiological differences, I focus here on broad features of phonological design. Earlier, I suggested that algebraic optimization defines human phonology. I thus proceed to ask whether it is also found in the learned vocal patterns of nonhumans.

As it turns out, neither of the two masters of the phonological mind—the algebraic or the phonetic—is uniquely human. There is ample evidence that productive algebraic rules are learnable by several species, both in natural communication of non-ape species [40,42] and laboratory settings [41,43,82,83]. Nonhuman vocal learners are likewise known to adapt their vocal patterns to phonetic patterns [84]. But while these two ingredients (the algebraic and phonetic) are not uniquely human, their conjunctive fusion—i.e., algebraic optimization—has not been shown in other species (Box 3). To the extent this plausibly presents evidence of absence, rather than mere absence of evidence (a question awaiting future research), the lacuna would indicate that algebraic optimization does not spontaneously emerge from its ingredients. Such a result would suggest that this capacity is the product of human adaptation to form phonological patterns.

Indeed, precursors of the phonological mind are already detected in early development. Young infants and neonates are capable of learning phonological rules [37]; they are sensitive to various phonological primitives (features, syllables, and consonants versus vowels [18,51,52,85]) and restrict their co-occurrence in line with putatively universal phonological restrictions [14].

But as in other systems of core knowledge, core phonological knowledge not only defines our early linguistic instincts, but also offers scaffolds for cultural inventions later in life. The intimate link between phonology and reading is a case in point. Unlike phonology, reading is neither instinctive nor early, and its acquisition typically requires considerable effort. Yet, all known writing systems,
both conventional systems [86] and those spontaneously invented by children [87,88], encode phonological units (e.g., features, segments, consonants, and syllables). Moreover, skilled readers automatically decode phonological structure from print even when they read silently [89,90].

Further clues into the design of the phonological mind are offered by dyslexia. In view of the close links between the phonological system and reading, one might expect individuals with dyslexia to exhibit deficits in linguistic phonological competence. But the actual state of affairs is more nuanced. Although dyslexics do in fact show numerous auditory and phonetic deficits [91–93] along with difficulties in gaining conscious awareness of phonological structure in spoken language [94], their grammatical phonological competence is apparently intact [72,95,96]. The preservation of the phonological grammar in dyslexia demonstrates its resistance to perturbations affecting the auditory–phonetic interface, and underscores their separability. The resilience of the phonological grammar is in line with its view as an innate core biological system.

Concluding remarks
The universality of phonology, its reliance on algebraic rules that are shared across languages and are functionally grounded, its early onset, and its role in reading are consistent with the view of phonology as an algebraic system of core knowledge. Whether certain phonological rules are, indeed, universal or innate remains an open empirical question (Box 4). Rather than offering a definitive answer, my goal here is to raise this possibility, draw attention to relevant findings, and outline their broad implications for areas ranging from psycholinguistics and reading research to neuroscience, language evolution, and comparative animal studies. I hope that this discussion promotes closer scrutiny of the phonological mind.

Acknowledgments
This research was supported by grant NIDCD DC003277. I wish to thank Ray Jackendoff and Paul Smolensky for their comments on this manuscript. I also thank Alex Catullo for technical assistance.

References
17 Perfetti, C. (2011) Phonology is critical in reading: but a phonological deficit is not the only source of low reading skill. In Explaining Individual Differences in Reading: Theory and Evidence (Brady, S.A. et al., eds), pp. 163–171, Psychology Press
49 Erickson, J.E. et al. (2010) Sensing the coherence of biology in contrast to psychology: young children’s use of causal relations to distinguish two foundational domains. Child Dev. 81, 390–409
81 Fehér, O. et al. (2009) De novo establishment of wild-type song culture in the zebra finch. Nature 459, 564–568
93 Lehongre, K. et al. (2011) Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. Neuron 72, 1080–1090
95 Berent, I. et al. (2012) Dyslexia impairs speech recognition but can spare phonological competence. PLoS ONE 7, e44875