Phonological Acquisition: Recent Attainments and New Challenges*

Sharon Peperkamp

Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS - ENS - CNRS, Paris and Université de Paris 8

Key words

phonological acquisition

phonological theory

segments

phonotactics

phonemes

Abstract

Infants' phonological acquisition during the first 18 months of life has been studied within experimental psychology for some 30 years. Current research themes include statistical learning mechanisms, early lexical development, and models of phonetic category perception. So far, linguistic theories have hardly been taken into account. These theories are based upon the assumption that there is a common core of innate phonological knowledge across speakers of all human languages, and they contain detailed proposals concerning phonological representations and the derivations by which abstract underlying forms are mapped onto concrete surface forms. It remains to be investigated experimentally if there is innate phonological knowledge and how the language-specific phonological grammar is acquired.

In the present article, the contributions to this special issue are introduced, and an attempt is made to bridge the gap between phonological theory and experimental psychology. In particular, some recent experimental work is considered in the light of phonological theories and new research avenues are sketched. What might be innate, what needs to be acquired, and how this acquisition might take place are questions that are addressed with respect to several aspects of phonological knowledge, specifically segmental representations, phonotactics, phonological processes, and the architecture of the phonological grammar.

1 Introduction

The astonishing speed and ease with which infants acquire their native language has long amazed both experts and laymen. Without any explicit instruction, children begin to utter their first words by the end of their first year of life, and by around three years, they have mastered important aspects of the grammar of their language and have developed a considerable lexicon. A now predominant research enterprise hypothesizes that first language acquisition begins bottom-up from the acoustic

* *Acknowledgments:* For valuable comments and discussion, I would like to thank Nick Clements, Emmanuel Dupoux, and Jessica Maye. Financial support from the CNRS ("Cognition et traitement de l'information," nº CTI02_15) is gratefully acknowledged.

Address for correspondence: Sharon Peperkamp, Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole Normale Supérieure, 46 Rue d'Ulm, 75005 Paris, France; e-mail: <sharon.peperkamp@ens.fr>.

signal; the acquisition of the sound structure of the native language thus lies at its basis. The collection of articles in the present issue deals with aspects of this early phonological acquisition, as well as with the interaction between phonological acquisition and early word learning, and with L2 phonological acquisition.

Some 30 years ago, experimental psychologists began to study infants' phonological development in perception and the way in which this allows bootstrapping into higher levels of language acquisition. The pioneering work of Peter Eimas, Peter Jusczyk, Patricia Kuhl, Jacques Mehler, and Janet Werker has thus brought phonological acquisition into the center of developmental psychology. To summarize briefly, it has been shown that during the first year of life, infants' perception of linguistic sound structures is shaped by the phonological properties of their native language. Infants begin by identifying global rhythmic and prosodic aspects of utterances in their language, before focusing on smaller units. They then acquire the repertoire of vowels and consonants of their language, start segmenting words, and acquire word-level properties such as metrical structure and phonotactics.

Set within the framework of experimental psychology, the majority of articles in the present issue are concerned with infants' developing phonology, as evidenced by changes in their perceptual capacities. Specifically, the contribution by Pierrehumbert, as well as those by Anderson, White, and Morgan, Best and McRoberts, and Bosch and Sebastián-Gallés deal with the acquisition of native phonetic categories. Fennell and Werker subsequently focus on the phonological representation of early words, as does Swingley. The final article, by Kingston, considers the acquisition of non-native segmental categories in adults. My aim is not to discuss the new research questions that arise out of these articles, but rather to look beyond the topics currently investigated and to focus on a broader theme: the integration of linguistic theory into developmental psychology.

Since the founding work of Chomsky and Hallé (1968), linguists have made detailed proposals concerning phonological representations and the derivations by which abstract underlying forms are mapped onto concrete surface forms. Most aspects of these proposals have not been looked at in experimental work on phonological acquisition. Furthermore, within the framework of generative grammar, it is assumed that there is a common core of phonological knowledge across speakers of all human languages. This common core is typically supposed to be innate (even though there is no a priori reason that universals could not emerge during the course of acquisition). Whether innate knowledge is used or not during phonological acquisition also remains to be investigated experimentally. After providing an overview of the contributions to this issue, I will thus consider phonological theory from the perspective of first language acquisition and focus on two questions: what exactly needs to be learned; and how could this learning take place, knowing that infants' initial input consists of the acoustic signal. Considering experimental approaches to phonological acquisition, I will review some work in the light of phonological theories and sketch new research avenues.

2 Overview

In the first article following this introduction, **Pierrehumbert** considers the vast amount of language-specific phonetic detail that speakers master and that hence needs to be learned by young children. This detail concerns both segmental and suprasegmental structure. Different languages do not employ the same boundaries to categorize sounds on a continuum, nor do they use the same cues and cue weightings in the acoustic realization of segmental and suprasegmental features. The same ambiguity is found within individual languages: A single sound may belong to different phonemes, depending upon the prosodic context in which it occurs, and segments likewise have multiple acoustic realizations. Pierrehumbert proposes that the segmental categories that are learned by infants are defined by positional variants, that is, segments couched within a segmental and prosodic context, rather than by phonemes or phones. The initiation of these categories is extremely rapid in infants, but refinement continues well into childhood and may last for certain aspects till over 12 years of age. Within an exemplar-based theory of speech processing, Pierrehumbert proposes that sound categories are acquired bottom-up by means of statistical learning procedures. Refinement of categories relies on two mechanisms: external feedback that is provided by a perception-production loop involving both the learning child and the speakers in his/her environment, and internal phonological feedback provided by type statistics over the lexicon. Finally, Pierrehumbert illustrates that the set of human languages is a subset of the set of mathematically possible languages. The proposed acquisition scenario with its major role for bottom-up learning indeed imposes constraints upon the structure of human languages that mathematically possible languages do not need to obey.

The issue of statistical learning mechanisms is taken up by **Anderson, Morgan, and White**, who consider the relative order in which infants acquire segmental categories. It is well known that newborns discriminate segmental contrasts regardless of whether they are used in their native language, but that at the end of the first year of life they have difficulties discriminating non-native contrasts, just like adults do. This decline in discrimination performance is generally interpreted as a sign of native category formation. Anderson et al. examine if input frequency plays a role in the acquisition of native categories. Specifically, they compare English-learning infants' capabilities to discriminate two non-native stop contrasts, one involving voiceless coronals and one involving voiceless dorsals. They show that at 8.5 months of age, performance on the coronal contrast but not on the dorsal one declines, suggesting that at this age, infants have acquired the voiceless coronal stop category of their native language but not the corresponding dorsal one. Given that the former is more frequent than the latter in English, these results are in accordance with the hypothesis that more frequent segmental categories are acquired before less frequent ones. Anderson et al. propose to account for the frequency effect within an attractor model. They also raise theoretical arguments against an alternative account, according to which universally unmarked segments such as coronals would be acquired before more marked ones. Their results do not allow them to distinguish the two accounts, since—as is often the case—markedness and frequency go hand in hand. Research with contrasts that do not show the same correlation between markedness and frequency is thus needed to settle this issue empirically.

Two more articles deal with the acquisition of segmental categories during the first year of life. Like Anderson et al., **Best and McRoberts** examine differences in the developmental patterns of non-native contrast discrimination. In a discrimination task with 6–8 and 10–12 month-old infants, they find that 10–12 month-olds show the same amount of decline on three isiZulu contrasts that are assimilated differently by adults. What these contrasts have in common is that they involve segments with the same place of articulation but that differ with respect to their laryngeal properties. Best and McRoberts account for the results by supplementing Best's Perceptual Assimilation Model (Best, 1994; Best, McRoberts, & Sithole, 1988) with the hypothesis that during the first year of life, infants are sensitive only to the primary articulatory organ of speech sounds (Browman & Goldstein, in preparation). Given that the differences in place of articulation between the sounds in the isiZulu contrasts correspond to differences in primary articulatory organ, this explains why 10–12 month-olds assimilate the isiZulu contrasts to a single category; the presence of articulatory differences among the contrasts that are responsible for different assimilation patterns in adults, then, are not relevant for infants. A further prediction of the revised model is that infants should maintain discrimination of non-native contrasts between segments produced with different primary articulatory organs. This prediction is borne out by a second experiment, in which infants show no decline on a Tigrinya contrast between consonants involving different primary articulators.

Bosch and Sebastián-Gallés examine native category formation in bilingual infants. They test infants from three linguistic environments at 4, 8, and 12 months of age on a vowel contrast, [e-ɛ], that is present in Catalan but not in Spanish. As expected, monolingual Catalan infants discriminate this contrast well at all ages, while Spanish monolinguals show a decline at 8 and 12 months. As to Catalan-Spanish bilinguals, they also show a decline at eight months but regain discrimination at 12 months. These results show that mere exposure to a contrast is not enough to maintain the capacity to perceive it. In their account of these facts, Bosch and Sebastián-Gallés make appeal to statistical learning as well. They show that in a bilingual environment, tokens of the three vowels exhibit a unimodal distribution, due to the fact that Spanish [e] falls in between the Catalan categories [e] and [ɛ] and has a high frequency compared to both Catalan vowels. Given findings by Maye, Werker, and Gerken (2002) that infants are sensitive to the distribution of tokens along a continuum, this might explain why eight-month-old infants who are exposed to both languages fail to discriminate the Catalan contrast: the monomodal distribution provides evidence for the existence of a single category. Bosch and Sebastián-Gallés tentatively propose that both quantity of exposure and lexical knowledge might play a role in teasing apart the Spanish category from the Catalan ones and hence the recovery of discrimination performance at 12 months.

The next two articles deal with phonological representations of early words. Previous studies on early word learning suggested that infants initially represent words with little phonological detail (see, for instance, Hallé & De Boysson-Bardies, 1996; Shvachkin, 1948/1973; and Stager & Werker, 1997). The articles by Fennell and Werker and Swingley contest this view and argue that, on the contrary, early word representations are phonologically precise encodings. **Fennell and Werker** use the Switch task to show that 14-months-olds are sensitive to phonological features in the representation of words they already know. They argue that the lack of such sensitivity found earlier with the same task when infants were tested on new words (Stager & Werker, 1997) is due to a processing overload. This latter hypothesis can be put to test by testing infants on new words using a less demanding task, such as a visual fixation procedure (Swingley & Aslin, 2000, 2002).

In his contribution, **Swingley** likewise investigates the phonological representation of familiar words. Building on previous work that provided evidence for precise encoding of early words (Swingley & Aslin, 2000, 2002), he shows that 19-month-old Dutch infants recognize words better when they are pronounced correctly then when one consonant is substituted by another one. The present study presents two novelties with respect to the previous ones. First, it is shown that the mispronunciation effect is obtained with word-medial substitutions as well as with word-initial ones, eliminating the possibility that only word-beginnings are encoded with phonological detail. Second, Swingley examines the hypothesis put forward by Charles-Luce and Luce (1990) that refinement of phonological representations is induced by the presence of phonological neighbors. Contrary to this view, Swingley and Aslin (2002) already found that the mispronunciation effect is independent of the size of infants' receptive vocabulary. In the present article, a broader definition of neighbor is introduced, according to which any chunk of speech that children might store as a potential word form and that differs from the target in a single segment counts as a neighbor. Swingley then compares substitutions with a frequent consonant, [d], to those with a very infrequent one, [g]. Under the broad definition of neighbor, several mispronunciations with [d] constitute neighbors of the tested words, but none with [g] does, due to the rarity of this consonant in Dutch. The mispronunciation effect crucially appears to be as large with [g] as with [d]. Moreover, some of the test words had no phonological neighbors at all, whether with [d] and [g] or with any other consonant. These words still yielded a mispronunciation effect. From these data, Swingley infers that phonological neighbors — even under the broader definition — play no role in the development of early lexical representations.

The last article in this issue deals with the learning of non-native sound categories by adults. **Kingston** reports on experiments in which American English listeners are trained to categorize German nonlow vowels, differing on the dimensions [back], [high] and [tense]. He shows, firstly, that for some contrasts different instances yield different results. For instance, subjects perform much better on the tense:lax distinction realized on the pair [ø-œ] than on the pair [y-ʏ]. Likewise, the presence of irrelevant context and speaker variation hinders subjects in their abilities to generalize to novel tokens for some contrasts but not for others. Finally, when subjects are asked to classify vowels into natural classes rather than into single phoneme categories, their performances again differ according to the defining features of these natural classes. Together, the results suggest that in some cases, subjects learn to extract abstract phonological features, while in others, they learn concrete sets of exemplars. Kingston argues that exemplar-based learning, but not feature-based learning, is most compatible with Kuhl's Native Language Magnet model (Kuhl, 1991, 2000), in which sound categories are prototypes that occupy a certain location in an acoustico-phonetic space. As to the observed differences between the three dimensions, he suggests that they might be due to the fact that features differ in the range of phonetic properties by which they are instantiated across all possible minimal pairs, that is, in how polymorphous they are. Extracting abstract features that are highly polymorphous is more difficult than extracting those that are less polymorphous. Finally, having focused on adult speech perception, Kingston argues that a fuller understanding of infants' acquisition of sound categories likewise requires the comparison of multiple contrasts realized in multiple pairs of segments and varying irrelevantly across speakers and context.

3 Experimental psychology and theoretical linguistics: Towards an integrated account of phonological acquisition

The collection of articles in this issue gives an indication of current research themes: statistical learning mechanisms, early lexical development, and models of phonetic category perception. As to statistical learning mechanisms, this topic has recently received much attention, not only in the realm of the formation of native segmental categories as in the present issue (see, for instance, Maye et al., 2002), but also with respect to word segmentation and higher-level grammatical regularities (see Gómez & Gerken, 2000, for an overview of research). It has thus been shown that statistical properties of speech can be picked up by infants. It remains to be investigated, though, if statistical learning is sufficient to acquire the grammar of one's native language. Generative linguists have raised the hypothesis that human languages share properties that, together, constitute Universal Grammar. Universal properties can be implications of the type "if a language has A then it also has B" or absolute constraints of the type "every language has C" and "no language has D." Whether or not there is innate knowledge that infants use during the course of acquisition is largely an open question.

Five of the seven articles in this issue deal with the acquisition of segmental categories, while the remaining two consider the representation of segments in early words. Hence, segments are at the heart of the matter. This is probably not a coincidence: more broadly, experimental research on speech perception has always shown a strong focus on segments, whether in adults or in infants. Since the nineteen-eighties, though, evidence has also been gathered as to when (and to a lesser extent how) other aspects of phonological knowledge are processed by adults and acquired during infancy. These aspects notably include phonotactics, metrical structure, and word segmentation (see Jusczyk, 1997, for a recent overview concerning infant development). It should be noted that all of these aspects concern surface properties of phonological structure. However, native speakers' phonological competence also includes knowledge about how surface structures are derived from abstract underlying structures. It is not enough to acquire the segmental repertoire, the possible combinations of these segments in syllables and words, and the stress patterns of words; infants should also learn how these surface properties arise when morphemes are strung together into words and words into phrases. That is, infants need to acquire their language's phonological grammar.

So far, experimental research on phonological acquisition has thus hardly taken linguistic theories into account, neither with respect to the presence of innate phonological knowledge nor with respect to the acquisition of the phonological grammar.[1] In the remaining part of this article, I consider how experimental approaches to phonological acquisition could incorporate phonological theory. In line with the tendency in this issue, and for the sake of simplicity, I consider segmental phonology only. I begin by discussing segmental representations, then turn to phonotactics and phonological processes that segments are involved in, and end by looking at the architecture of the phonological grammar that derives surface forms from underlying forms. For all of these aspects, I examine what might be innate, what needs to be learned, and how this acquisition might take place.

3.1
Segments and phoneme inventories

One can pose many questions concerning the presence of innate phonological knowledge and the way in which segmental representations develop in first language acquisition. To start with the most basic one, do infants come equipped with knowledge about the limited numbers of categories that are distinguished in the IPA-table, and is all they have to do learn which of these categories are present in their language? Along these lines, Kuhl (2000) argues that at birth, infants categorize the vowel space such as to separate the vowels of all human languages; a statistical analysis of the acoustic signal in their native language then leads them to wipe out unnecessary boundaries. At least some of the innately present boundaries are thought to be "natural," in the sense of being determined by the auditory system, since nonhuman animals are sensitive to them as well. By contrast, in her contribution to the present issue, Pierrehumbert (2003) argues that given the amount of language-specificity, segmental (and, for that matter, prosodic) categories must be acquired bottom-up in the absence of universal category boundaries.

Segments are not atomic units. From an articulatory phonetic point of view, they

[1] The reverse is equally true: Linguistic approaches to phonological acquisition have been little concerned with experimental evidence regarding early phonological development. In particular, there has been a strong focus on children's early productions, which— following Jakobson (1941)—have been taken to reflect children's phonological knowledge (see Bernardt & Stemberger, 1998, and references cited therein). These studies thus sidestep the fact that before uttering their first words, infants acquire important parts of their native language's phonology. Moreover, the learning algorithms that have been developed so far within linguistic theory (see, among others, Boersma, & Hayes, 2001; Dresher, 1999a; Dresher & Kaye, 1990; Hayes, in press; Pulleyblank & Turkel, 2000; Tesar & Smolensky, 1998) share a feature that makes them rather unrealistic. That is, they take as their input individual words, whereas infants likely acquire many aspects of phonological structure before their lexicon of word forms is big enough to use word-based algorithms. Discussion of these issues, however, is outside of the scope of the present article.

are described in terms of three main dimensions: place of articulation, manner of articulation, and voicing. According to the motor theory of speech perception, humans have innate knowledge concerning the articulatory properties of sounds (Liberman & Mattingly, 1985); these three-dimensional phonetic representations could thus be present at birth. In phonological theory, manner, place, and voicing of segments are further defined by one or more phonetic features each, most of which are binary (Chomsky & Hallé, 1968; Sagey, 1986). Natural classes, that is, sets of segments that share certain articulatory properties, can thus be defined as feature bundles. For instance, the class of voiced obstruents is defined as [–sonorant, +voice]. It is generally assumed that phonetic features are universal too, and, therefore, might be present innately. For instance, in his quantal theory of speech, Stevens (1972, 1989) argues that phonological features are determined by natural boundaries that arise from the nonmonotonic relationships between acoustic and articulatory parameters by which speech sounds are defined. Pulleyblank (2001), however, questions the universality of phonological features on the basis of both typological arguments and arguments from cross-linguistic phonological processes. Phonetic features, he argues, have concrete perceptual or articulatory correlates that may differ across languages; they should thus emerge during the course of acquisition.

The role of subsegmental representations during the acquisition of segmental categories has recently been studied in both adults and infants. Recall that in his contribution to the present issue, Kingston shows that American English adults can learn to extract certain abstract phonological features from German. As to first language acquisition, Maye and Weiss (2003) tested eight-month-old American English infants' capacities to discriminate a Hindi contrast involving a prevoiced and an unaspirated stop after being habituated to a continuum of tokens that lie in between the two endpoints. They showed that after a short exposure to a bimodal but not to a monomodal distribution of tokens, infants discriminated the endpoints. Crucially, they also found that infants generalized their discrimination capacity to the same VOT-contrast realized at a different place of articulation. That is, when trained on alveolar stops, they discriminated velar stops, and the other way round. Maye and Weiss argue that these results are evidence that infants extract featural properties of input speech. They are in line with research by Jusczyk, Goodman, and Baumann (1999), who found that nine-month-olds are sensitive to the manner of articulation of word-initial consonants. That is, when presented with two types of word lists, infants showed a listening preference for those lists in which the initial consonants of the words had the same manner of articulation. Note, however, that neither of these experiments shows that infants represent segments as bundles of phonetic features. Rather, the conclusion that can be drawn is that infants are sensitive to basic articulatory and/or acoustic properties of segments.

To summarize, segments are made up of distinctive features, which reflect articulatory properties of speech sounds.[2] These features, furthermore, exhibit two types of dependency relationships among each other. The first one is phonetic: some

[2] The earliest proposals for distinctive features were actually based on acoustic properties of speech sounds (Jakobson, Fant, & Hallé, 1952).

features are relevant only to make a distinction within a category defined by another feature. For instance, [apical] is relevant only for coronal segments, which are themselves defined by the feature [coronal]. Therefore, [apical] is a dependent of [coronal]. The second type of dependency relationship is phonological: certain features function together in phonological rules and constraints. For instance, in consonantal place assimilation (as in English *impossible* and *incomplete*, the latter with the cluster [ŋk]), the major place features [coronal], [labial], and [dorsal] and their dependents spread from one consonant to another, to the exclusion of any nonplace features, such as [sonorant] and [voice]. Hence, these major place features are interdependent. Within linguistic theory, it is proposed that the phonetic features by which segments are composed are organized hierarchically in a feature tree. This proposal goes under the name of feature geometry (Clements, 1985; Clements & Hume, 1995; Hallé, 1995; Sagey, 1986).[3] It is generally assumed that there is a universal feature hierarchy, although no consensus has been reached concerning its exact nature. In order to account for the above mentioned facts, [apical] should be a daughter node of [coronal], and [coronal], [labial], and [dorsal] should be bundled together under a place node in this universal hierarchy. So far, no experimental research concerning the existence of a set of hierarchically organized features has been conducted, neither in adult speech perception nor in infant development.

Linguistic theory is not just concerned with the surface speech stream segmented into consonants and vowels. Besides this concrete surface level, it recognizes an abstract underlying level of representation. Language-specific phonological rules and alternations constitute the mapping of underlying representations onto surface representations. As far as segmental representations are concerned, the underlying level contains abstract phonemes rather than concrete segments. Segments, then, are the phonetic realizations of phonemes, and a given phoneme can have multiple phonetic realizations, depending upon both its segmental and prosodic context. Hence, the set of segments present in a language is larger than its set of phonemes, and the mapping of underlying phonemes onto surface segments is language-specific.

The representation of phonemes at the underlying level is less detailed than the representation of segments at the surface level we have just seen. For instance, only those features are typically argued to be present in phonemes that are needed to encode all the lexically distinctive contrasts in a particular language. The number of features that are thus needed is correlated with the size of the language's phoneme inventory: A three-vowel system, for instance, can be described with only two features, while a seven-vowel system needs four. However, the question as to exactly which features are needed is less straightforward, because there are often multiple ways in which a given inventory can be described without redundancies. For instance, consider the three-vowel system /i,u,a/, with the feature specifications as in Table 1.

[3] An alternative theory of the internal structure of segments is offered by dependency phonology (Anderson & Ewen, 1987; van der Hulst, 1989).

TABLE 1

Feature specifications in a three-vowel system

	high	*low*	*back*	*round*
/i/	+	–	–	–
/u/	+	–	+	+
/a/	–	+	+	–

This system can be described with any of the pairs of features [high] and [back], [high] and [round], [low] and [back], [low] and [round], and [round] and [back]. The choice for a set of features is often determined by the phonological behavior of the segments. For the case at hand, if [u] and [a] pattern together for some phonological process to the exclusion of [i], then [back] should be present, whereas if [i] and [a] pattern together to the exclusion of [u], then [round] should be present. In the absence of such phonological evidence, however, there is an indeterminacy problem.

Often, some version of underspecification is assumed, according to which not only redundant features are absent, but redundant feature specifications can be left blank as well.[4] For instance, in many languages, voicing is contrastive in obstruents but not in sonorants. For purposes of the description of the consonant inventory, sonorants therefore do not need to have a specification for the feature [voice] at the underlying level but can receive their voicing value by a late redundancy rule. Whether or not a feature is left unspecified underlyingly is often motivated by phonological behavior of segments. For instance, consider languages in which voicing is not contrastive in sonorants and that, moreover, contain a voice assimilation process. In most of these languages, only obstruents, not sonorants, trigger assimilation. If voice assimilation is defined as spreading of the feature [voice] onto an adjacent segment, this is evidence that in these languages, sonorants are unspecified for [voice] at the underlying level. In the remaining languages in which assimilation is triggered by both sonorants and obstruents, by contrast, [voice] should be specified in sonorants underlyingly, despite the fact that voicing is not lexically contrastive in sonorants. There are often multiple ways in which, given a segment inventory and a set of nonredundant features, the feature specifications can be filled in without creating redundancies, and a language's phonological processes do not always provide evidence in favor of one of the options. Hence, underspecification raises a second indeterminacy problem.

A solution to both indeterminacy problems is proposed by Clements (2001). He distinguishes two abstract levels of representation, a lexical level and a phonological level, with the latter being derived from the former. As to the lexical level, Clements proposes that representations are determined by a universal scale of feature accessibility. More accessible features take priority over less accessible ones as far as the

[4] There are currently several different proposals concerning underspecification; see Steriade (1995) for an overview and Dresher (1999b) for discussion in the realm of acquisition.

specification of segments in an inventory is concerned, and they are hence less prone to omission in redundancy situations. For instance, consider a language with the consonantal inventory /p,t,k,m,n/. The highly accessible feature [sonorant] distinguishes /m/ and /n/ from the other consonants. The feature [nasal], which is less accessible, is therefore redundant and omitted from their representations. Likewise, in languages with both voiced and voiceless obstruents but only voiced sonorants, obstruents are specified for [voice] but sonorants are not. In fact, the voicing distinction that exists between, for instance, /p/ and /m/ is subsumed under the distinction made by the more accessible feature [sonorant]. Thus, a feature can be completely absent from the lexical level (as is the case for [nasal] in the first example), or it can be specified in some segments while left unspecified in others (as is the case for [voice] in the second example).

The universal accessibility hierarchy is based upon typological considerations: Features that are recurrent across languages are highly accessible, whereas rare features are placed low in the accessibility hierarchy. Given that all languages have both consonants and vowels, [consonantal] is the most accessible feature. Next in the hierarchy are [coronal] and [sonorant], respectively. At the bottom we find features such as [constricted], needed in languages that distinguish between plain and glottalized consonants. At first sight, the feature accessibility hierarchy is a likely candidate for being innately present, since infants do not have access to typological facts. However, the bare existence of strong typological tendencies suggests that the universal accessibility hierarchy might be grounded in acoustic properties of speech sounds. An alternative possibility, then, is that the acoustic signal contains information that allows infants to construct this hierarchy.

Turning now to the phonological level, Clements argues that it is derived from the lexical one by the addition of those features and feature specifications that are necessary to account for a language's phonology. For instance, suppose that the three-vowel inventory /i,a,u/ is described at the lexical level by [high] and [round]. The same two features are carried over to the phonological level. If in a language with this inventory the back vowels [u] and [a] pattern together for some phonological process to the exclusion of [i], then [back] is added at this level. Similarly, consider languages with a voicing contrast in obstruents but not in sonorants. If sonorants pattern together with voiced obstruents in that they trigger voice assimilation, then the specification for [voice] will be added to the sonorants. As to the hierarchical organization of features, Clements (2001) proposes that at the phonological level, a given language only has the amount of structure in the feature tree that is necessary to account for its phonology. Crucially, all constituents in a language's featural representations must be constituents of the universal feature geometry. For instance, many languages need both [posterior] and [strident] to differentiate multiple coronal consonants in their segment inventory. Some of these languages have an assimilation process that involves [posterior] but not [strident], showing evidence that [posterior] must be projected on its own tier. The reverse situation, assimilation of [strident] without [posterior], never occurs. In the universal feature geometry, [posterior] therefore is a daughter of [strident]. Representational economy is illustrated by those languages in which [strident] and [posterior] always act together; here, [posterior] is attached to the same node as [strident], thus yielding a simpler structure than in the universal geometry.

This short overview shows that representations at the phonetic surface are richer than those at the underlying level(s) in several respects. As said before, in phonological theory, surface representations are derived from underlying ones; structure is thus added during the derivation of surface representations. Infants are in the reverse situation: Given that they acquire their language from the surface speech stream, they should rather prune redundant information from the representations as they acquire the underlying structures of their native phonology. In particular, the following acquisition scenario suggests itself. First, infants acquire the set of segments used in their language by a bottom-up statistical analysis of the speech signal. They then build minimal feature representations of these segments, working their way down the feature accessibility hierarchy, and construct the corresponding feature trees, in line with the universal feature geometry. From these surface representations, they subsequently build underlying phonological representations, in which the feature tree only contains the amount of structure needed to account for the language's phonology and in which phonologically redundant features and feature specifications are absent. Finally, they construct even more abstract lexical representations, in which only lexically contrastive distinctions are encoded and in which the feature tree has no structure at all.[5]

Many questions remain open in this scenario. Most importantly, the language-specific representations at the underlying levels depend upon both the phoneme inventory and the phonological processes present in the language. These aspects should thus be acquired prior to the construction of underlying representations. Below, I examine how infants might do this. Before turning to this topic, however, I consider an additional aspect of segmental representations at the surface level, that is, phonotactic constraints.

3.2
Phonotactic constraints

The way in which segments can be combined into syllables and words is restricted by language-specific phonotactic constraints. These constraints are static, in the sense that they capture generalizations over the lexicon that do not result from phonological rules and alternations. Well-known examples from English are the prohibitions on syllable-final [h] and word-initial [ŋ], on the onset clusters [sr] and [ʃl] (as opposed to legal [ʃr] and [sl]), and on lax vowels in stressed open syllables. Adults are sensitive to the phonotactic constraints of their language (Hallé, Segui, Frauenfelder, & Meunier, 1998; Hay, Pierrehumbert, & Beckman, in press; Massaro & Cohen, 1983; Pitt, 1998), and there is evidence that infants acquire the phonotactic constraints of their language at around nine months of age (Jusczyk, Luce, & Charles-Luce, 1994).

---

5 The scenario sketched here is the reverse of the one commonly proposed in child phonology. According to these latter proposals, infants initially assume all sounds to be variants of a single phoneme. By making binary distinctions between sounds, following an order that is echoed by Clements' feature accessibility hierarchy, they build up an inventory in which all distinctive sounds are differentiated (Jakobson, 1941; Jakobson & Hallé, 1956). A recent version of such a scenario that incorporates feature geometry and underspecification is proposed by Rice and Avery (1995).

The nature of phonotactic constraints has long been a topic in phonological theory. It has been shown that, crosslinguistically, constraints on the presence of segments in certain syllabic positions show a robust asymmetry between onsets and codas. The set of consonants allowed in coda position is typically smaller, if anything, than that in onset position. Many languages indeed display limitations upon the consonants that can be found in syllable codas. For instance, Japanese only allows a placeless nasal consonant and the first half of a geminate consonant in the syllable coda (Shibatani, 1990). Similar restrictions on the set of consonants that may appear in syllable onsets are much rarer. Coda conditions, moreover, often concern natural classes of segments. For instance, Beijing Chinese does not allow obstruents in coda position (Blevins, 1995). Finally, coda constraints typically favor segments that are prone to assimilate (such as coronal and placeless consonants), as well as segments with a high sonority. It would be truly surprising if a language were found in which syllable codas can contain dorsals but not coronals, or obstruents but not sonorants.

As far as acquisition is concerned, it remains to be investigated how infants acquire phonotactic constraints, and, more specifically, if the acquisition mechanism is sensitive to the naturalness of distributional regularities. Recently, Chambers, Onishi and Fisher (2003) argued that phonotactic constraints are learned at the level of individual segments. They exposed 16.5-month-old American infants to CVC-words in which the onset and coda consonants constituted distinct classes. For one group of subjects, /b,k,m,t,f/ appeared in the onset and /p,g,n,tʃ,s/ in the coda, while for another group it was the reverse. During testing, infants discriminated between novel items that obeyed the distributional regularities they had been exposed to and those that did not, thus showing sensitivity to phonotactic constraints that are not present in their native language. Notice that the two classes of consonants in this experiment do not form natural classes; that is, they cannot be differentiated by a single phonetic feature or set of features. Although it is true that individual segments can be engaged in this type of regularities, the set of regularities used by Chambers et al. do not seem to reflect a natural language. To the best of my knowledge, there is indeed no language in which as many as 10 segments occur in onsets or codas exclusively, with the onset and coda groups crosscutting several natural classes. It would be interesting to see if infants are even more sensitive to this type of distributional regularity if the consonants that appear in onsets and codas, respectively, form natural classes.

A first step towards exploring the role of natural classes in the acquisition of phonotactic constraints is taken by Saffran and Thiessen (2003). They exposed nine-month old infants to disyllabic CVC.CVC words that displayed certain regularities. In one experiment, the consonants that occurred in onsets and codas, respectively, formed complementary natural classes. That is, syllables began with a voiced stop and ended with a voiceless one for one group of subjects while the reverse was true for another group. Using a word segmentation task, Saffran and Thiessen showed that infants found it easier to segment novel words out of a continuous speech stream that were in accordance with the pattern to which they had been exposed than those that were not. In another experiment, they showed that the same did not hold if, during the exposure phase, the consonants that appeared in onsets and codas, respectively, did not form natural classes. The set of consonants used in both experiments being identical, these results suggest that phonotactic regularities that involve natural classes of segments are easier to acquire than those that do not.

Given the differences in both the experimental design and the age of the infants, it is hard to compare the results of Saffran and Thiessen (2003) to those of Chambers et al. (2003). Both techniques, however, allow us to further address the question concerning the acquisition of natural versus unnatural regularities.

3.3 *Phonological processes*

Languages contain many surface regularities that involve phonological processes and that therefore do not generally fall under the heading of phonotactics. For instance, in French, [e] can only occur in open syllables; this constraint gives rise to alternations such as *premier* [pʀəmje] 'first$_{\text{MASC}}$'—*première* [pʀəmjɛʀ] 'first$_{\text{FEM}}$' and *céder* [sede] 'to give up'—*cède* [sɛd] '(I) give up'. Hence, /e/ turns into [ɛ] in closed syllables. Likewise, in French, obstruent clusters agree in voicing. This holds both within and across words, and a rule of regressive voice assimilation applies whenever a disagreeing cluster arises when words are strung together in a phrase, as in *robe sale* [ʀɔpsal] 'dirty dress' (cf. *robe jaune* [ʀɔbʒon] 'yellow dress'). In a similar vein, we can formulate phonological generalizations that are observable at the phonetic surface and that concern the distribution of allophones. For instance, in American English, /t/ is realized as [ɾ] if it occurs in between a sonorant sound and an unstressed vowel, as in *wai[ɾ]er* (cf. *wai[t]ress*), and *wai[ɾ] a minute* (cf. *wai[t] for me*). The surface generalization is that [t] does not occur between a sonorant segment and an unstressed vowel.

Given the evidence concerning infants' developing sensitivity to static phonotactic constraints during the first year of life (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993), it seems likely that they similarly acquire the surface true generalizations that underlie phonological rules and alternations. One piece of evidence in this direction is provided by Jusczyk, Hohne, and Bauman (1999), who showed that 10.5-month-old English-learning infants can rely on the allophonic differences between *nitrates* and *night rates* in a word segmentation task, suggesting that they have acquired at least some aspects of the distribution of the various allophones of English /t/ and /ɪ/.[6] When and how infants acquire the precise nature of rules and alternations, however, is still an open question. Consider the case of French voice assimilation. Becoming sensitive to the fact that obstruent clusters agree in voicing does not suffice; infants should also learn that disagreeing clusters are eliminated by a rule of regressive voice assimilation (rather than by progressive voice assimilation, consonant deletion, or some other phonological process). It is this latter piece of knowledge that they need in order to infer, for instance, that [ʀɔp] in *robe sale* is an instance of the same word as [ʀɔb] in *robe jaune*. Similarly, American English learning infants should not only learn the distributions of [t] and [ɾ]; they should also learn that the two segments are realizations of a single phoneme. Again, it is this knowledge

6 Most importantly, in *nitrates*, but not in *night rates*, /t/ is aspirated and /ɪ/ is devoiced.

that will allow them to infer that [weɪɾ] in *wait a second* is the same word as [weɪt] in *wait for me*.

Perception experiments show that adults are sensitive to these types of generalizations: they treat allophonic contrasts, such as the one between [ɾ] and [t] in English, differently from phonemic ones (Jaeger, 1980; Lahiri & Marslen-Wilson, 1991; Pegg & Werker, 1997; Peperkamp, Pettinato, & Dupoux, 2003; Whalen, Best, & Irwin, 1997), and they undo the effects of assimilation rules for purposes of lexical access (Coenen, Zwitserlood, and Bölte, 2001; Darcy, Ramus, Christophe, Kinzler, & Dupoux, submitted; Gaskell & Marslen-Wilson, 1996, 1998). These data show that knowledge about phonological rules and alternations form part of native speakers' perceptual competence, and warrants the question as to when and how this knowledge is acquired.

According to a recent proposal by Steriade (2001), the type of phonological process that underlies a certain surface-true generalization is determined by a universal perceptual distance metric. Specifically, phonological processes are minimal from a perceptual point of view, and all languages that disallow a given structure apply the same process to avoid this structure. For instance, Steriade argues that across languages, the absence of disagreeing obstruent clusters is always due to regressive voice assimilation, rather than to progressive voice assimilation, consonant deletion, or some other phonological process, and that this is so because the outcome of regressive voicing is perceptually closer to the disagreeing cluster in the input than the outcome of any other process would be. Consequently, infants who have acquired a surface-true generalization would be able to infer the underlying phonological process by inspecting their own perception. Although it is certainly true that in the majority of languages, the absence of obstruent clusters that disagree in voicing is due to regressive voice assimilation, this is not the case in all languages. For instance, Dutch has a productive rule of progressive voice assimilation that devoices fricatives after voiceless obstruents (Booij, 1995). Other surface-true generalizations are likewise due to more than one phonological process. To give a well-known example, many languages do not allow for clusters consisting of a nasal consonant followed by a voiceless obstruent. A range of processes is found cross-linguistically to avoid such clusters at the surface, including obstruent voicing, coalescence, nasal deletion, and denasalization (Pater, 1999). Hence, a language's phonological processes cannot be acquired on the basis of a perceptual distance metric.

One might want to argue that acquiring phonological rules relies upon prior lexical acquisition. For instance, knowing that the phrases *robe jaune* and *robe sale* both refer to the same type of object allows French-learning infants to infer that [ʀɔb] and [ʀɔp] are instances of the same word. Peperkamp and Dupoux (2002), however, propose learning mechanisms based on distributional cues that are available to prelexical infants. They argue that knowledge of phonological processes facilitates lexical acquisition, since it reduces the list of word forms that should be mapped onto a meaning. Whether or not infants learn about the phonological rules in their language before having a lexicon is, of course, an empirical question. Here, I review the prelexical learning algorithms of Peperkamp and Dupoux and discuss several complicating factors.

Let us first consider allophonic rules. Once infants can represent speech as a sequence of consonants and vowels, they should discover which segments are involved in allophonic contrasts, that is, are realizations of a single phoneme. In other words, they should reduce the inventory of segmental categories to the inventory of abstract phoneme categories. This problem is far from trivial, since two segments that are realizations of a single phoneme — hence form an allophonic contrast — in one language can be realizations of two different phonemes — hence form a phonemic contrast — in another language. For instance, in French, voicing contrasts are phonemic in obstruents and allophonic in sonorants (Dell, 1973); by contrast, in Burmese, voicing is phonemic in sonorants (Okell, 1969) and in Korean, it is allophonic in obstruents (Kim, 1990). Now, a property that distinguishes segments in allophonic contrasts from those in phonemic ones is that they have complementary distributions. For instance, recall that in American English, [ɾ] is found between sonorant segments and unstressed vowels, whereas [t] is excluded in those positions. If infants detect these complementary distributions, they can infer that [ɾ] and [t] represent the same abstract phoneme. So far, the possible use of a distributional analysis by language learners has been tested only in adults, with positive results (Almeida, 2003; Peperkamp et al., 2003). In these experiments, which used two different artificial language learning paradigms, French adults were shown to be sensitive to complementary distributions for the purposes of the construction of phoneme categories. In order to further test the acquisition algorithm, the same type of experiment is now being carried out with infants.

Experiments with artificial language learning paradigms sidestep several complicating factors found in natural languages that are worth consideration. Many complications concern the nature of allophonic rules. For instance, some segments are allophones of more than one phoneme, as is the case of American English [ɾ] that can represent both an underlying / t / and an underlying / d /; therefore, the distribution of [ɾ] is complementary to that of both [t] and [d]. The reverse case, a phoneme with multiple allophones, also occurs. English / t /, for instance, can surface as any of the following consonants: [t, ɾ, tʰ, cʰ, t̚], all of which have nonoverlapping distributions. Complementary distributions can thus involve more than two segments. Moreover, allophonic rules often target natural classes of segments. For instance, aspiration in English involves all three voiceless stops / p,t,k /; consequently, not only the pairs [pʰ-p], [tʰ-t] and [kʰ-k] display complementary distributions, but so do [tʰ-p] and all the other pairs involving an aspirated and a plain stop. Allophonic rules typically apply within a prosodic domain, which creates a further complication. For instance, Greek has a rule of allophonic / s /-voicing that applies before voiced consonants within the intonation phrase (Nespor & Vogel, 1986); consequently, [s] and [z] have complementary distributions within but not across intonation phrases. In order to find complementary distributions of pairs of segments, infants should thus look for empty intersections in their distributions within prosodic domains, without there being a priori restrictions on the size of the domain in which rules apply. Phonological rule domains can indeed be as small as the syllable (as for English / l /-darkening) and as large as the utterance (as for English flapping) (Nespor & Vogel, 1986). It should also be noted that allophonic rules can be optional or gradient, in the sense that the likeliness that they apply decreases in larger prosodic domains.

All of the considerations above show that it can be hard to find the complementary distributions that are induced by allophonic rules. Conversely, languages can display complementary distributions of segments that linguists would not consider to be the result of an allophonic rule. For instance, English [h] and [ŋ] have complementary distributions within the domain of the prosodic word, since [h] occurs before stressed vowels only and [ŋ] is excluded from that position. Yet, phonologists do not generally consider [h] and [ŋ] to be realizations of a single phoneme, and hence they do not postulate an allophonic rule. The same holds for [œ] and [ɥ] in French, with the former occurring in closed syllables only — hence followed by a consonant — and the latter occurring before vowels only. Constraints on the distribution of segments can thus interfere with the learning algorithm. Likewise, segments with a low frequency might be engaged in spurious complementary distributions even in the absence of phonological restrictions on their distributions.

Cross-linguistically, phonological rules — whether allophonic or not — appear to conform to two intuitions. First, the output of phonological rules stays close to the input; for instance, there is no rule that maps an underlying /m/ onto a surface [k], which differs on all possible phonetic dimensions. Second, the conditioning context of rules shares certain properties with the output; for instance, languages may have a rule that maps /t/ onto [p] before labials and into [k] before velars, but there is no language in which /t/ is mapped onto [p] before, say, high vowels. The theory of feature geometry (Clements, 1985; Clements & Hume, 1995; Hallé, 1995; Sagey, 1986) limits the class of possible rules by ordering features hierarchically in a feature tree as seen above, and by stipulating a limited number of possible operations. Specifically, all phonological processes are defined in terms of either spreading, delinking, insertion, or deletion of constituents of this tree. Allophonic relationships between [h] and [ŋ] in English and between [œ] and [ɥ] in French are thus excluded. It remains to be investigated if infants use this type of knowledge in order to eliminate candidate rules that are excluded by the theory of feature geometry. The artificial language-learning paradigms used by Peperkamp et al. (2003) and Almeida (2003) allow us to address this question, since they can compare the acquisition of possible rules to that of impossible ones.[7]

Turning now to the case of nonallophonic rules, Peperkamp and Dupoux (2002) show that they involve complementary distributions of complete word forms rather than of individual segments. Nonallophonic rules are neutralizing; that is, their outputs involve segments that occur in lexical contrasts. Consider again French voice assimilation, according to which obstruents assimilate in voicing to a following obstruent. I tentatively take the domain of this rule to be the phonological phrase. The final [p] of assimilated [ʀɔp] in *robe sale* occurs also in nonassimilated forms, such as *top* [tɔp] ‘pip’. In contexts other than before obstruents, both [b] and [p] occur, as in the isolated words *robe* and *top*; hence [b] and [p] do not have complementary

7 It is important to note that “in accordance with Universal Grammar” is not equivalent to “natural” in the sense of recurrent across languages. Indeed, phonological processes are not necessarily natural (Anderson, 1981). One might expect that infants have the least difficulties acquiring natural rules, followed by unnatural and impossible rules, respectively.

distributions. However, the word forms [ʀɔb] and [ʀɔp] do have complementary distributions. Indeed, [ʀɔp] occurs before voiceless obstruents within the same phonological phrase only, whereas [ʀɔb] occurs in all other contexts. Peperkamp and Dupoux argue that once infants build a repertoire of word forms, they can carry out a distributional analysis of word forms and infer that [ʀɔb] and [ʀɔp] are two forms of the same word.

As noted by Peperkamp and Dupoux, the proposed learning algorithm fails in one case. To see this, consider the minimal pair *rab* [ʀab] 'extra' – *rap* [ʀap] 'rap'. Here, the assimilated form of the former word is homophonous to the nonassimilated form of the latter word. The two phonetic forms of the word *rab*, then, do not have complementary distributions. Indeed, in contexts other than before obstruents (such as phrase-finally), both [ʀab] and [ʀap] occur. In order for the algorithm to be useful, the number of minimal pairs of this type should be small enough for infants to consider them as noise.

An altogether different way to look at nonallophonic rules is worth consideration. It is often argued that truly neutralizing rules do not exist (see Manaster Ramer, 1996, for a review of this literature). For the French case at hand, it is possible that the final [p] of assimilated *robe* is not completely homophonous to the final [p] of *top* and that the two instances of /p/ form distinct categories across tokens. The "neutralizing" voice assimilation rule would thus behave like an allophonic rule, in the sense of creating complementary distributions among segments. Hence, the algorithm proposed for allophonic rules would be applicable in this case, too.

A final complication for the acquisition of phonological processes concerns rule interaction. The overwhelming majority of languages contain multiple phonological rules, which can exhibit complicated interactions. For instance, in French, vowels are lengthened before voiced consonants. Hence, the vowel in *robe* is longer than that in *top*. As far as I know, the interaction between voice assimilation and this allophonic lengthening has not been studied. Consider once more *robe sale*, that presents the context for both rules. The application of voice assimilation is independent of that of vowel lengthening, but the reverse is not true, since lengthening critically depends upon the voicing feature in the following consonant. There are thus two possibilities: either voice assimilation blocks lengthening, due to the fact that it destroys the context for lengthening, or both rules apply. The latter case is the more interesting one, since it yields a surface form with a long vowel before a voiceless consonant. As a consequence, long and short vowels do not have complementary distributions at the surface. Allophonic lengthening is thus opaque, making its acquisition more difficult than as would follow from the algorithm of Peperkamp and Dupoux (2002). However, once this allophonic rule is acquired, the acquisition of voice assimilation is facilitated, since the presence of a lengthened vowel before a voiceless consonant now provides a cue that voice assimilation has taken place. Kiparksy (1973) already raised the hypothesis that surface opacity makes acquisition particularly hard. For the algorithm laid out above, this is so because the discovery of a surface true generalization is the first step in the acquisition of phonological processes.[8]

8 For more discussion on opacity and the problems it poses for acquisition, see Dresher (1999b).

To sum up, it is proposed that infants begin by acquiring surface true generalizations by a bottom-up distributional analysis, much the same way in which they acquire static phonotactic constraints. Given evidence that infants can segment the speech stream into clauses at seven months (Hirsh-Pasek et al., 1987), and into phrases and words at nine months (Jusczyk et al., 1992) and 11 months (Myers et al., 1996), respectively, surface true generalizations that hold within a large prosodic domain could be acquired before those that hold within smaller domains. A comparative analysis of the distribution of individual segments then allows infants to find complementary distributions and hence construct a set of hypothetical allophonic rules. From this set, they might select the true allophonic rules of their language by eliminating those that are ruled out in Universal Grammar. Finally, it is not until infants have built a reasonably large lexicon of word forms that they can acquire nonallophonic rules. One problem with this scenario was already pointed out above: a rule's surface-true generalization can be obscured, due to interaction with another rule. A second potential problem concerns the interaction between the discovery of surface true generalizations for allophonic rules on the one hand and the acquisition of segmentation of prosodic units on the other hand. In fact, whereas it is proposed here that the acquisition of allophones depends upon segmentation into prosodic units, it might be argued, conversely, that allophones provide cues for the parsing of continuous speech into prosodic units. Note, however, that prosodic boundaries are signaled not only by allophonic cues but also by suprasegmental cues such as final lengthening. A solution out of this circularity, then, might be that infants first acquire the segmentation into prosodic units on the basis of suprasegmental cues, which allows them to subsequently acquire the allophonic rules.

3.4
The architecture of the grammar

A final topic that I would like to discuss concerns the architecture of the grammar that deals with the derivation of surface forms from underlying forms. There are currently two classes of theories, one based on rules and the other one based on constraints. In rule-based theories, surface forms are derived from underlying forms by a system of ordered rewrite rules. In these theories, the absence of any rules at the initial state is uncontroversial: Infants have to acquire the rules that are present in their language, rather than "forgetting" those that are not. At the initial state, their grammar is thus the simplest possible one, in which underlying forms are mapped as such onto surface forms without intervening rules. During acquisition, infants use innate knowledge concerning the set of possible rules as defined by UG to discover the rules of their language.

In constraint-based theories, by contrast, the question of the nature of the initial state is less straightforward. In these theories, surface forms are derived from underlying forms by means of a system of ranked and violable constraints. In the most well-known version, Optimality Theory (Prince & Smolensky, 1993), there are two types of constraints: markedness constraints, that militate against (universally) marked output structures, and faithfulness constraints, that require output forms to be maximally similar to their corresponding inputs. How an input presenting some marked structure surfaces depends upon the relative marking of the relevant markedness and faithfulness constraints. By way of illustration, let us consider consonant clusters. Homorganic clusters, that is, clusters with a single place of articulation, are universally less marked than nonhomorganic ones. This follows from typological observations such as the following: There are languages that disallow nonhomorganic clusters, but there are no languages that disallow homorganic ones. What will happen to an input presenting a nonhomorganic cluster in a given language, for instance if a morpheme ending in a coronal consonant is concatenated to a morpheme beginning with a labial consonant? If the markedness constraint against such clusters outranks faithfulness, then markedness will be respected and the input will be transformed into an output form with a homorganic cluster (or no cluster at all). If, by contrast, faithfulness outranks markedness, then the underlying cluster will surface as such.

Turning now to the question as to the initial state of the grammar within Optimality Theory, there are three issues to be discussed. The first one concerns the origin of the set of markedness and faithfulness constraints. All constraints are generally assumed to be universal and, moreover, innate. Notable exceptions are Boersma (1998) and Hayes (1999). For instance, Hayes (1999) proposes that markedness constraints are emergent from experience in articulation and perception. In his proposal, infants construct candidate constraints randomly, based on innate knowledge of phonological features. They then assess these constraints for their degree of phonetic grounding. A phonetically grounded constraint is one that bans sound structures that are hard from an articulatory point of view. Among the candidate constraints, only the phonetically-grounded ones are maintained.

The second issue concerns the ranking of markedness constraints. Following markedness theory as developed first in structural linguistics (Trubetzkoy, 1939), certain classes of markedness constraints are assumed to show a fixed, internal, ranking that is the same for all languages (Prince & Smolensky, 1993). These so-called harmonic rankings, then, are innate and do not allow for reranking during acquisition. For instance, it is often argued that coronal is the unmarked place of articulation (see, among others, Paradis, & Prunet, 1991). The constraint that bans coronals is therefore universally ranked below the ones that ban consonants at other places of articulation. Markedness, however, reflects typological tendencies, not universals: Individual languages can deviate from this harmonic ranking and have an unmarked place of articulation other than coronal. In such cases, an *ad-hoc* constraint is needed on top of the hierarchy with the effect of undoing the harmonic ranking. For this reason, the use of harmonic scales has been criticized by Hume and Tserdanelis (2003), who argue that typological tendencies fall outside the domain of phonological theory.[9] For language-learning infants, this means that they have no preconceived idea as to which place of articulation is most likely the unmarked one in their language. The same conclusion is reached by Anderson et al. (2003) in the present issue on the basis of experimental findings. Recall that they found the decline in discrimination of non-native coronal contrasts to precede that of other contrasts. Anderson et al. consider the possibility that innate knowledge of the unmarked

[9] See also Hale and Reiss (1998), who make the same point concerning the fate of typological tendencies.

status of coronals might offer an explanation for the observed asymmetry during development, but conclude that, on the contrary, markedness cannot be at stake. They reason that if infants had such innate knowledge, they should maintain discrimination of non-native coronal contrasts *longer* than that of contrasts at other places or articulation, contrary to what was actually found. Indeed, one of the defining properties of the unmarked status of coronal is that across languages, coronal consonants occur in more varieties than consonants at other places of articulation; infants should thus be prepared to make more distinctions among coronals than among noncoronals and, hence, maintain discrimination of non-native coronal contrasts.[10]

The third and final issue regarding the initial state concerns the relative ranking of markedness and faithfulness constraints. Smolensky (1996) argues that all markedness constraints initially outrank all faithfulness constraints, yielding universally unmarked outputs only. As infants analyze their linguistic input, they subsequently demote those markedness constraints that ban the structures that are present at the surface in their language. This position is by far the most wide-spread, and can be considered an implementation of the hypothesis that infants' phonological development proceeds from universally unmarked structures to the marked structures present in their native language (Jakobson, 1941). With respect to consonant clusters, it implies that infants start out with a grammar that allows homorganic clusters only. If their language exhibits nonhomorganic clusters, they will demote the markedness constraint that demands homorganicity. The rationale behind this view of the initial state and the process of phonological acquisition is of the type *reductio ad absurdum*. That is, if, conversely, faithfulness initially dominated markedness, final grammars would be overly permissive. To see this, suppose a language that has only homorganic clusters. With faithfulness initially outranking markedness, infants would start out with a grammar that allows both homorganic and nonhomorganic clusters. Given that this initial grammar is in accordance with their native language, they would have no reason to promote the constraint that demands homorganicity; consequently, at the final state their grammar would still allow nonhomorganic clusters.

One might seriously question, though, if infants will leave open the possibility that their language permits nonhomorganic clusters after having analyzed many hours of continuous speech containing homorganic clusters only. That is, the persistent absence of certain marked structures might play a role during the course of acquisition. Taking this argument seriously, one might propose that at the initial state, faithfulness initially outranks markedness. This is indeed the position taken by Hale and Reiss (1998). As noted by these authors, this scenario is much closer to the one for rule-based grammars that everybody would agree upon, in that infants start

[10] Anderson et al. also argue that their results on the acquisition of segmental categories cannot be accounted for by learned knowledge of the unmarked status of coronals in English. In fact, this knowledge can arise from either a study of the consonant inventory or a comparison of the processes applying to consonants at various places of articulation; the examination of inventories and rules, however, requires the prior learning of segmental categories.

out with the simplest possible grammar, consisting of an identity mapping from underlying forms onto surface forms.

In a series of experiments, Davidson, Smolensky, and Jusczyk (in press) investigated the question of the relative ranking of markedness and faithfulness at the initial state. Using the Headturn Preference Procedure, they presented four and six-month-old infants with triads of the form [X]… [Y]… [Z], where [Z] was the concatenation of [X] and [Y]. In the crucial experiment, [X] and [Y] were chosen such that a marked structure was present at the boundary between these two elements. The concatenation was either faithful, thus exhibiting this marked structure, or unfaithful, yielding an unmarked structure. For instance, if [X] was [in] and [Y] was [pa], then Z was either [inpa] or [impa]. The former concatenation is completely faithful with respect to the presumed input /inpa/, but it is more marked than the latter, since it presents a nonhomorganic consonant cluster. In the latter form, a homorganic cluster is obtained by means of the well-known process of place assimilation. Given that until six-months of age, infants do not develop a sensitivity to the surface regularities of their language (Jusczyk et al., 1994), Davidson et al. assumed the infants' grammar to be in its initial state with respect to the regularities under investigation. The main result of the experiments was that infants listened longer to the unfaithful triads of the form [in]… [pa]… [impa] then to triads of the form with the markedness violation [in]… [pa]… [inpa]. Moreover, if markedness was held constant, infants listened longer to faithful ([im]… [pa]… [impa]) than to unfaithful ([im]… [pa]… [uŋkɚ]) triads, and if faithfulness was held constant, they preferred less marked ([im]… [pa]… [impa]) to more marked ([in]… [pa]… [inpa]) triads. Davidson et al. thus concluded that at the initial state, both markedness and faithfulness constraints are present, and that the former crucially outrank the latter.

Note, however, that the logic of the experiments relies upon the assumption that infants suppose [Z] to be the output of the concatenation of /X/ and /Y/. This assumption is questionable, especially in the case of triads of the form [im]… [pa]… [uŋkɚ], where [uŋkɚ] should thus correspond to underlying /impa/.[11] Moreover, given that in the Headturn Preference Procedure, infants sometimes show a familiarity preference and sometimes a novelty preference, it is virtually impossible to interpret the results in terms of their grammar. The conclusion that markedness outranks faithfulness at the initial state therefore seems unwarranted. Instead, the results could show that infants prefer to listen to words containing homorganic consonant clusters rather than nonhomorganic ones, and that they prefer triads that exhibit a certain homogeneity. The issue of the ranking of markedness and faithfulness at the initial state, then, remains an open question.

---

[11] To investigate the role of faithfulness constraints, one might alternatively compare faithful triads of the form [tim]… [po]… [timpo] and [po]… [tim]… [potim] to their unfaithful counterparts [tim]… [po]… [potim] and [po]… [tim]… [timpo], with the latter presenting a well-known process of metathesis.

4 Conclusion

Neither the discussion of the representation of segments and segmental processes in phonological theory nor that of experimental work concerning these aspects is meant to be exhaustive. For instance, I have not touched upon the intricacies of the interaction between morphology and phonology, as studied in the theory of lexical phonology (see, for instance, Mohanan, 1986). What I hope to have illustrated, though, is that what needs to be learned largely exceeds the topics that have been looked at so far in experimental approaches to phonological acquisition. Of course, adding suprasegments and the processes they are involved in will only further complicate the picture. I refer to Pierrehumbert's contribution to the present issue for discussion on this topic.

It is hoped that the present issue will stimulate research that aims at bridging the gap between phonological theory and experimental psychology. To this end, the research agenda should include, on the one hand, a focus on theoretical considerations concerning the learnability of phonological structure and the development of concrete algorithms that are based on the observation that infants acquire their language bottom-up; and, on the other hand, experimental explorations of phonological acquisition that take into account the fact that human languages have phonological systems whose complexity is tremendous, but limited by the possibilities of Universal Grammar.

References

ALMEIDA, D. (2003). *L'acquisition des catégories phonémiques.* Unpublished MA thesis, EHESS, Paris.

ANDERSON, J., MORGAN, J., & WHITE, K. (2003). (this issue). A statistical basis for speech sound discrimination. *Language and Speech*, **46**, 155–182.

ANDERSON, J. M., & EWEN, C. J. (1987). *Principles of dependency phonology.* Cambridge, U.K.: Cambridge University Press.

ANDERSON, S. R. (1981). Why phonology isn't natural. *Linguistic Inquiry*, **12**, 493–539.

BERNARDT, B., & STEMBERGER, J. (1998). *Handbook of phonological development.* New York: Academic Press.

BEST, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

BEST, C. T., & McROBERTS, G. W. (2003). (this issue). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, **46**, 183–216.

BEST, C. T., McROBERTS, G. W., & SITHOLE, N. M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 345–360.

BLEVINS, J. (1995). The syllable in phonological theory. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Oxford: Blackwell.

BOERSMA, P. (1998). *Functional phonology.* The Hague: Holland Academic Graphics.

BOERSMA, P., & HAYES, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, **32**, 45–86.

BOOIJ, G. E. (1995). *The phonology of dutch*. Oxford: Oxford University Press.

BOSCH, L., & SEBASTIÁN-GALLÉS, N. (2003). (this issue). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, **46**, 217–243.

BROWMAN, C. P., & GOLDSTEIN, L. (in preparation). *Articulatory phonology*. Cambridge, MA: MIT Press.

CHAMBERS, K. E., ONISHI, K. H., & FISHER, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, **87**, B69–B77.

CHARLES-LUCE, J., & LUCE, P. A. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language*, **17**, 205–215.

CHOMSKY, N., & HALLE, M. (1968). *The sound pattern of english*. New York: Harper & Row.

CLEMENTS, N. (1985). The geometry of phonological features. *Phonology Yearbook*, **2**, 225–252.

CLEMENTS, N. (2001). Representational economy in constraint-based phonology. In T. A. Hall (Ed.), *Distinctive feature theory* (pp. 71–146). Berlin: Mouton de Gruyter.

CLEMENTS, N., & HUME, E. (1995). The internal structure of speech sounds. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 245–306). Oxford: Blackwell.

COENEN, E., ZWITSERLOOD, P., & BÖLTE, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, **16**, 535.

DARCY, I., RAMUS, F., CHRISTOPHE, A., KINZLER, K., & DUPOUX, E. (submitted). Language specific effects in compensation for phonological variation.

DAVIDSON, L., SMOLENSKY, P., & JUSCZYK, P. (in press). The initial and final states: Theoretical implications and experimental explorations of richness of the base. In R. Kager & J. Pater & W. Zonneveld (Eds.), *Constraints in phonological acquisition*. Cambridge, U.K.: Cambridge University Press.

DELL, F. (1973). *Les règles et les sons*. Paris: Hermann.

DRESHER, E. (1999a). Charting the learning path: Cues to parameter setting. *Linguistic Inquiry*, **30**, 27–67.

DRESHER, E. (1999b). Child phonology, learnability, and phonological theory. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 299–346). San Diego: Academic Press.

DRESHER, E., & KAYE, J. (1990). A computational learning model for metrical phonology. *Cognition*, **34**, 137–195.

FENNELL, C., & WERKER, J. (2003). (this issue). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, **46**, 245–264.

GASKELL, M. G., & MARSLEN-WILSON, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 144–158.

GASKELL, M. G., & MARSLEN-WILSON, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 380–396.

GÓMEZ, R., & GERKEN, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, **4**, 178–186.

HALE, M., & REISS, C. (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, **29**, 656–683.

HALLE, M. (1995). Feature geometry and feature spreading. *Linguistic Inquiry*, **26**, 1–46.

HALLÉ, P. A., & BOYSSON-BARDIES, B. de (1996). The format of representation of recognized words in the infants' early receptive lexicon. *Infant Behavior and Development*, **19**, 463–481.

HALLÉ, P. A., SEGUI, J., FRAUENFELDER, U., & MEUNIER, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 592–608.

HAY, J. B., PIERREHUMBERT, J., & BECKMAN, M. (in press). Speech perception, well-formedness, and the statistics of the lexicon. In R. Ogden & J. Local & R. Temple (Eds.), *Phonetic interpretation. Papers in laboratory phonology VI.* Cambridge, U.K.: Cambridge University Press.

HAYES, B. (1999). Phonetically driven phonology: The role of Optimality Theory and inductive grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer, & K. Wheatly (Eds.), *Functionalism and formalism in linguistics* (Vol. 1: General papers, pp. 243–285). Amsterdam: John Benjamins.

HAYES, B. (in press). Phonological acquisition in Optimality Theory: The early stages. In R. Kager & J. Pater & W. Zonneveld (Eds.), *Constraints in phonological acquisition.* Cambridge, U.K.: Cambridge University Press.

HIRSH-PASEK, K., KEMLER NELSON, D. G., JUSCZYK, P. W., CASSIDY, K. W., DRUSS, B., & KENNEDY, L. (1987). Clauses are perceptual units for young infants. *Cognition*, **26**, 269–286.

HUME, E., & TSERDANELIS, G. (2003). Labial unmarkedness in Sri Lankan Portuguese Creole. *Phonology*, **19**, 441–458.

JAEGER, J. J. (1980). Testing the psychological reality of phonemes. *Language and Speech*, **23**, 233–253.

JAKOBSON, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell.

JAKOBSON, R., FANT, G., & HALLE, M. (1952). *Preliminaries to speech analysis.* Cambridge, MA: MIT Press.

JAKOBSON, R., & HALLE, M. (1956). *Fundamentals of language.* The Hague: Mouton.

JUSCZYK, P., FRIEDERICI, A., WESSELS, J., SVENKERUD, V., & JUSCZYK, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, **32**, 402–420.

JUSCZYK, P., GOODMAN, M. B., & BAUMANN, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language*, **40**, 62–82.

JUSCZYK, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.

JUSCZYK, P. W., HIRSH-PASEK, K., NELSON, D. G., KENNEDY, L. J., WOODWARD, A., & PIWOZ, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, **24**, 252–293.

JUSCZYK, P. W., HOHNE, E. A., & BAUMAN, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, **61**, 1465–1476.

JUSCZYK, P. W., LUCE, P. A., & CHARLES-LUCE, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, **33**, 630–645.

KIM, N.-K. (1990). Korean. In B. Comrie (Ed.), *The world's major languages* (pp. 881–898). Oxford: Oxford University Press.

KINGSTON, J. (2003). (this issue). Learning foreign vowels. *Language and Speech*, **46**, 295–349.

KIPARKSY, P. (1973). Abstractness, opacity, and global rules. In O. Fiujimura (Ed.), *Three dimensions of linguistic theory* (pp. 57–86). Tokyo: TEC.

KUHL, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, **50**, 93–107.

KUHL, P. K. (2000). Language, mind, and brain: Experience alters perception. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed.) (pp. 99–115). Cambridge, MA: MIT Press.

LAHIRI, A., & MARSLEN-WILSON, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, **38**, 245–294.

LIBERMAN, A. M., & MATTINGLY, I. (1985). The motor theory of speech perception revised. *Cognition*, **21**, 1–36.

MANASTER RAMER, A. (1996). A letter from an incompletely neutral phonologist. *Journal of Phonetics*, **24**, 477–489.

MASSARO, D. W., & COHEN, M. M. (1983). Phonological context in speech perception. *Perception & Psychophysics*, **34**, 338–348.

MAYE, J., & WEISS, D. (2003). Statistical cues facilitate infants' discimination of difficult phonetic contrasts. In B. Beachley & A. Brown & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (pp. 508–518). Sommerville, MA: Cascadilla Press.

MAYE, J., WERKER, J. F., & GERKEN, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**, B101–B111.

MOHANAN, K. P. (1986). *The theory of lexical phonology.* Dordrecht: Reidel.

MYERS, J., JUSCZYK, P. W., KEMLER NELSON, D. G., CHARLES-LUCE, J., WOODWARD, A. L., & HIRSH-PASEK, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language*, **23**, 1–30.

NESPOR, M., & VOGEL, I. (1986). *Prosodic phonology.* Dordrecht: Foris.

OKELL, J. (1969). *A reference grammar of colloquial Burmese.* Oxford: Oxford University Press.

PARADIS, C., & PRUNET, J.-F. (Eds.). (1991). *The special status of coronals: Internal and external evidence.* New York: Academic Press.

PATER, J. (1999). Austronesian nasal substitution and other NC? effects. In R. Kager & H. v. d. Hulst & W. Zonneveld (Eds.), *The Prosody-Morphology Interface* (pp. 310–343). Cambridge, U.K.: Cambridge University Press.

PEGG, J. E., & WERKER, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, **102**, 3742–3753.

PEPERKAMP, S., & DUPOUX, E. (2002). Coping with phonological variation in early lexical acquisition. In I. Lasser (Ed.), *The process of language acquisition* (pp. 359–385). Frankfurt: Peter Lang.

PEPERKAMP, S., PETTINATO, M., & DUPOUX, E. (2003). Allophonic variation and the acquisition of phoneme categories. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (Vol. Volume 2, pp. 650–661). Sommerville, MA: Cascadilla Press.

PIERREHUMBERT, J. (2003). (this issue). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, **46**, 115–154.

PITT, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics*, **60**, 941–951.

PRINCE, A., & SMOLENSKY, P. (1993). *Optimality Theory. Constraint interaction in generative grammar.* Unpublished manuscript, Baltimore, MD.

PULLEYBLANK, D. (2001). *Defining features and constraints in terms of complex systems. Is UG too complex?* Handout of a talk presented at the Workshop on Early Phonological Acquisition, held in Carry-le-Rouet, October 6–8, 2001.

PULLEYBLANK, D., & TURKEL, W. (2000). Learning phonology: Genetic algorithms and Yoruba tongue-root harmony. In J. Dekkers & F. v. d. Leeuw & J. v. d. Weijer (Eds.), *Optimality Theory. Phonology, syntax, and acquisition* (pp. 554–591). Oxford: Oxford University Press.

RICE, K., & AVERY, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 23–42). Hillsdale, NJ: Lawrence Erlbaum Associates.

SAFFRAN, J. R., & THIESSEN, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, **39**, 484–494.

SAGEY, E. (1986). *The representations of features and relations in nonlinear phonology*. Unpublished Doctoral dissertation, MIT, Cambridge, MA.

SHIBATANI, M. (1990). Japanese. In B. Comrie (Ed.), *The world's major languages* (pp. 855–880). Oxford: Oxford University Press.

SHVACHKIN, N. K. (1948/1973). The development of phonemic speech perception in early childhood. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 91–127). New York: Holt, Rinehart, and Winston.

SMOLENSKY, P. (1996). *The initial state and "richness of the base" in Optimality Theory* (Technical Report). Baltimore, MD: Cognitive Science Department, Johns Hopkins University.

STAGER, C. L., & WERKER, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, **388**, 381–382.

STERIADE, D. (1995). Underspecification and markedness. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 114–174). Oxford: Blackwell.

STERIADE, D. (2001). *The phonology of perceptibility effects: The P-map and its consequences for constraint organization*. Unpublished manuscript.

STEVENS, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In P. B. Denes & E. E. D. Jr., (Eds.), *Human communication. A Unified View* (pp. 51–66). New York: McGraw Hill.

STEVENS, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, **17**, 3–46.

SWINGLEY, D. (2003). (this issue). Phonetic detail in the developing lexicon. *Language and Speech*, **46**, 265–294.

SWINGLEY, D., & ASLIN, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, **76**, 147–166.

SWINGLEY, D., & ASLIN, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, **13**, 480–484.

TESAR, B., & SMOLENSKY, P. (1998). Learnability in Optimality Theory. *Linguistic Inquiry*, **29**, 229–268.

TRUBETZKOY, N. S. (1939). Grundzüge der Phonologie. *Travaux du Cercle Linguistique de Prague*, **7**.

Van Der HULST, H. (1989). Atoms of segmental structure. *Phonology*, **6**, 253–284.

WHALEN, D. H., BEST, C. T., & IRWIN, J. R. (1997). Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics*, **25**, 501–528.