ELSEVIER

Rapid acquisition of phonological alternations by infants

Katherine S. White [a,*], Sharon Peperkamp [b], Cecilia Kirk [c], James L. Morgan [a]

[a] Brown University, Providence, RI, USA
[b] Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS, DEC-ENS, CNRS), Université de Paris VIII, and Institut Universitaire de France, France
[c] University of Canterbury, Christchurch, New Zealand

Abstract

We explore whether infants can learn novel phonological alternations on the basis of distributional information. In Experiment 1, two groups of 12-month-old infants were familiarized with artificial languages whose distributional properties exhibited either stop or fricative voicing alternations. At test, infants in the two exposure groups had different preferences for novel sequences involving voiced and voiceless stops and fricatives, suggesting that each group had internalized a different familiarization alternation. In Experiment 2, 8.5-month-olds exhibited the same patterns of preference. In Experiments 3 and 4, we investigated whether infants' preferences were driven solely by preferences for sequences of high transitional probability. Although 8.5-month-olds in Experiment 3 were sensitive to the relative probabilities of sequences in the familiarization stimuli, only 12-month-olds in Experiment 4 showed evidence of having grouped alternating segments into a single functional category. Taken together, these results suggest a developmental trajectory for the acquisition of phonological alternations using distributional cues in the input.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Infant speech perception; Statistical learning; Phonological alternations

* Corresponding author. Address: Department of Brain and Cognitive Sciences, University of Rochester, Meliora Hall, Rochester, NY 14627, USA. Tel.: +1 585 275 7192.
E-mail address: kwhite@bcs.rochester.edu (K.S. White).

1. Introduction

The first year of life is marked by rapidly developing knowledge of native language structure. Infants' changing speech perception abilities reflect acquisition of segmental categories (i.e., vowels and consonants; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Polka & Werker, 1994; Werker & Tees, 1984) and phonotactic structure (Jusczyk, Luce, & Charles-Luce, 1994), as well as the beginnings of fluent speech segmentation (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Jusczyk & Aslin, 1995). This progression is likely driven in large part by statistical analyses of the input speech (Anderson, Morgan, & White, 2003; Chambers, Onishi, & Fisher, 2003; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996).

In this article, we address another, under-explored, aspect of phonological development: infants' acquisition of phonological processes operating in their language. When segments, morphemes, or words are juxtaposed, phonological processes may alter their surface forms, introducing alternations among perceptually distinct speech sounds (Trubetzkoy, 1939/1969). Discovering the relations among alternating segments is critical for language processing; according to many conceptions of language representation, listeners must be able to "undo" these processes to relate surface forms to underlying representations (Gaskell & Marslen-Wilson, 1996; Lahiri & Marslen-Wilson, 1991).

The speech signal is inherently variable. Tokens of words and morphemes are physically distinct, due to both non-linguistic changes (e.g., talker, speech rate) and modifications induced by phonological processes. Despite this variability, listeners must be able to recognize the equivalence of tokens in different contexts. Previous work has suggested that the ability to detect the equivalence of speech sounds (Kuhl, 1983) and words (Houston & Jusczyk, 2000; Singh, Morgan, & White, 2004) altered by non-linguistic variation emerges during the first year. Although accounts of how listeners deal with non-linguistic phonetic variability differ, ultimately, this type of variability is irrelevant for morphological or lexical identity. Phonological variation, by contrast, caused by the sorts of phonological processes discussed in this paper, is potentially much more problematic. This is because segments that alternate at one level may be contrastive at another level, as will be explained below. Therefore, the nature of the phonological context in which a segment occurs is critical. Very little research has focused on infants' ability to cope with phonological variation.

We distinguish among three types of phonological processes. One type of process, allophony, relates phonologically conditioned phonetic variants of the same underlying phonemic category. For example, in English, $[p^h]$ and $[p]$ are allophones of the phoneme /p/: $[p^h]$ occurs syllable-initially; $[p]$ surfaces elsewhere. Thus, $[p^h]$ and $[p]$ *alternate* as a function of phonological context; the phonetic difference between them is not contrastive (i.e., does not produce a semantic contrast). Importantly, adults have difficulty distinguishing among allophones of the same phoneme; they perceive alternating allophones as more similar than phones in different phonemic categories, even when acoustic distance is equated (Whalen, Best, & Irwin, 1997). Moreover, language-specific effects of phonemic status have been found both behaviorally and neurally (Kazanina, Phillips, & Idsardi, 2006).

Using an oddball paradigm and magnetoencephalographic (MEG) recordings, Kazanina et al. revealed a mismatch negativity response (a neural marker for change detection) when two sounds [d] and [t] were drawn from separate phonemic categories in the listener's language, but not when they were allophones of the same phonemic category. Because only phonemic contrasts produce a change in meaning, learners must determine which phonetic distinctions in their language are phonemic and which are allophonic.

Allomorphic processes relate variants of the same morpheme. Allomorphs surface as phonologically conditioned variants of the same underlying morpheme. For example, the English plural/possessive/3rd person singular morpheme surfaces as [s] following most voiceless segments and as [z] following most voiced segments (the exception being voiced and voiceless sibilants, which are followed by [əz]). It would be impossible to acquire many aspects of morphology, such as plural and tense markings, without implicit recognition that such phonetically distinct manifestations are functionally equivalent. Unlike allophonic variation, the distinction between allomorphs, such as [s] and [z], is contrastive at another level: both /s/ and /z/ are phonemes in English. In some contexts, the difference between these two sounds signals a difference in meaning: "sip" and "zip", for example, are different words. When occurring as morphemes in appropriate phonological contexts, however, this distinction disappears: "cat" and "dog" are different stems, but suffixing either [s] to the former or [z] to the latter modifies the meaning in precisely the same way.

This duality, in which a phonetic distinction is contrastive in most but not all contexts, is also found in allolexic (sandhi) processes that alter the phonetic forms of words.[1] For instance, in English, coronal consonants like [t] and [n] can acquire place features from the following consonant. These coronal consonants can thus sound more like [p] and [m], respectively, before labials ("swee[p] boy", "te[m] pairs"), and more like [k] and [N] before velars ("swee[k] girl", "te[N] cars").

There is evidence to suggest that adults are able to map alternating forms onto a single lexical item (Coenen, Zwitserlood, & Bölte, 2001; Gaskell & Marslen-Wilson, 1996, 1998). In word-monitoring tasks, adults successfully identify tokens that have been modified by the surrounding phonological context. For example, adults identify the first part of "brow[m] bear" as "brown", even though the final segment has been modified by its proximity to the following labial segment.

Adults have two converging sources of information that may enable them to successfully detect alternating forms. First, they have lexical knowledge. When an adult hears the sequence "browm bear", the lexical entry for "brown" will be activated; "browm" is not a known word. Second, adults have knowledge of phonological

---

[1] We use the term allolexic to describe processes that introduce variability in the surface forms of words, by analogy with allophonic and allomorphic processes, which introduce variability at the level of phones and morphemes, respectively. Allolexic processes, like allomorphic processes, may also alter suprasegmental properties of words, avoiding stress clash ("fourTEEN tromBONES" vs. "FOURteen TRUMpets"; Selkirk, 1984) or breaking up sequences of identical tones (Clements, 1978). In this article, we confine our discussion to processes affecting the segmental properties of words.

alternations. Knowing that coronal nasals may be assimilated before labial conso-
nants will further help the lexical-processing system to resolve the item in favor of
"brown". Though compensation of this sort may be due in part to available phonetic
cues (i.e., listeners' ability to attribute features of the perhaps incompletely assimi-
lated form to the original segment, or feature parsing; Gow, 2003), language-specific
phonological knowledge appears to play a role as well. Listeners show a lack of com-
pensation for non-native assimilation patterns. For example, English speakers fail to
compensate for voicing assimilation, which occurs in French but not English; by
contrast, French speakers fail to compensate for place assimilation, which occurs
in English but not French (Darcy, Ramus, Christophe, Kinzler, & Dupoux, in press).

Knowledge of allophonic, allomorphic, and allolexic alternations would similarly,
in principle, be useful for language learners. However, little work exists on infants'
acquisition of phonological alternations, even though these types of alternations
are potentially much more disruptive to learners.

There is evidence that 10.5-month-olds are *sensitive* to allophonic variation and
can use this variation to segment words (Jusczyk, Hohne, & Bauman, 1999; Mattys
& Jusczyk, 2001). While these results indicate that infants have learned something
about the distributional properties of allophonic variation in English (e.g., aspirated
phones occur syllable-initially), they do not address whether infants understand the
relationship between allophones of the same phoneme. Pegg and Werker (1997)
tested adults' and infants' perception of two acoustically similar phones from two
different phonemic categories of English (syllable-initial [d] and a voiceless unaspi-
rated coronal stop (the [t] following an [s])). Adults were able to discriminate the
two sounds, though not at the high level typical of native phonemic contrasts. Young
infants were able to discriminate the sounds, but by 12 months their discrimination
was poor. These results show that acoustic similarity and phonological status inter-
act. Although the two sounds are drawn from different phonemes, they occur in non-
overlapping phonological contexts and thus are not contrastive; moreover, the two
sounds are highly similar acoustically. However, this study leaves open the question
of how the perception of acoustically distinct phones from the *same* phonological
category changes as a function of language exposure. Nor is there any work address-
ing whether infants understand the relationship between allomorphs of the same
morpheme, or allolexes of the same word.

Infants demonstrate sensitivity to bound morphemes, like the English plural/pos-
sessive/3rd person singular – s marker, in the second year (Soderstrom, White, Con-
well, & Morgan, 2007). Other work has focused on when older infants and children
understand the function of such markings (Kouider, Halberda, Wood, & Carey,
2006). However, these studies do not explicitly explore whether infants can equate
phonetically different forms of the same functional morpheme. For example, it could
be the case that children understand the plural function of [s] while still failing to
map the same plural function onto [z]. Likewise, there is little or no research on
how infants cope with phonological processes that alter the surface forms of adjacent
words.

Infants in the early stages of building a lexicon cannot know *a priori* which types
of variation are contrastive (e.g., phonemic) and which types are not (e.g., allo-

phonic, allomorphic, allolexic, non-linguistic). That is, infants without a lexical entry for "brown" cannot use lexical knowledge to determine whether a modified form maps onto a lexical entry. In fact, very young infants appear to weight non-linguistic and lexically irrelevant variation across tokens quite heavily, at least when learning new words (Bortfeld & Morgan, submitted for publication; Houston & Jusczyk, 2000; Singh et al., 2004, in press). This work has demonstrated that the ability to recognize the equivalence of word tokens that differ acoustically along lexically irrelevant dimensions appears to emerge in the latter half of the first year. Phonological variation is potentially more problematic because, as noted earlier, these types of alternations can involve segments that are sometimes contrastive.

How might phonological alternations be acquired? One possibility is that learners rely on semantic information to determine the relationship between alternating sounds. For example, if there are no minimal pairs distinguished by a phonetic contrast, learners may infer that the contrast is an allophonic one. Thus the absence of minimal pairs differing only in the [pʰ]–[p] distinction may lead English learners to hypothesize that these two sounds form an allophonic contrast. On the other hand, when phonetic differences, such as between [p] and [b], lead to a distinction in meaning, sounds will be retained in separate phoneme categories. Semantic knowledge may also be used more directly. A learner who observes that phonetically different forms are applied to the same referent (e.g., as when "bottle" is produced with a word-medial glottal stop or flap) may hypothesize that there is a relationship between these variants and group them together. This type of knowledge may be necessary for inferring the relationship between forms in free variation. However, generalization would require observation of these mappings across many different referent-label pairs. Therefore, a generalized understanding of phonological alternations derived using semantic strategies could only occur relatively late in development.

We hypothesize that pre-lexical infants may be able to make use of a different type of information to learn phonological alternations. That is, infants may be able to overcome the limitations of an immature lexicon by extracting information about phonological regularities from the speech signal. Specifically, learners may be able to use distributional information to determine the relationship between alternating sounds (Peperkamp & Dupoux, 2002). Several types of phonological alternations are characterized by complementary, non-overlapping distributions. Thus, learners may be able to group similar sounds occurring in complementary distribution. For example, if [pʰ] occurs only syllable-initially, but the phonetically similar [p] occurs elsewhere, learners may be able to infer that these sounds should be grouped into a single category. Complementary distributions of this sort can concern surface segments, surface morpheme forms, or surface word forms. Although some types of phonological alternations are not learnable in this way (e.g., free allophonic variation), infants equipped with sensitivity to such distributional patterns could break into the system during the first year of life.

We further hypothesize that phonetic similarity plays a role in constraining this type of learning process. It would be surprising to find that English listeners group [h] (which never occurs in syllable-final position) and [N] (which occurs exclusively in syllable-final position) into the same category. For adults, similarity is likely deter-

mined by both phonetic and phonological factors. Judgments of coarse-grained phonetic similarity (e.g., [h] and [N]) probably differ little as a function of language background. However, it is likely that perception of more fine-grained similarity differs across languages: adult native speakers of Thai, a language with a three-way phonemic VOT contrast, perceive the difference between [p] and [pʰ] to be as great as the difference between [p] and [b]; English speakers, by contrast, hear [p] and [pʰ] as more similar than [p] and [b]. Less influence of phonological knowledge should be observed in young infants. For infants who have not yet constructed functional categories, [p] likely sounds as similar to [pʰ] as to [b].

Currently, there are no data that directly address whether infants are capable of analyzing patterns of complementary distribution. However, recent work has demonstrated the potential power of statistical learning to explain many aspects of early language knowledge (Anderson et al., 2003; Chambers et al., 2003; Maye et al., 2002; Saffran et al., 1996). For example, corpus analyses have revealed differences in the frequency with which speech sounds occur; these frequency differences influence the trajectory of phonetic category learning, such that infants learn categories for more frequent speech sounds first (Anderson et al., 2003). A recent corpus study by Werker et al. (2007) has demonstrated that distributional information is available in the input to distinguish one vs. two phonetic categories along a single acoustic dimension. Behavioral experiments have found that infants' sensitivity to this type of distributional information affects phonetic category formation (Maye et al., 2002). Thus, not only is the relevant distributional information available in the input, but infants have been shown to use it. Complementing this behavioral and corpus work, computational work by Peperkamp, Le Calvez, Nadal, and Dupoux (2006a) has demonstrated that a statistical algorithm that searches for segments in complementary distribution robustly discovers allophonic distributions.

In the present work, we investigate whether infants can use distributional information alone to infer that two normally contrastive segments are related by a phonological process. We familiarized infants with artificial languages whose distributional properties exhibited one of two phonological alternations, involving stop or fricative devoicing.[2] It is important to note that, unlike static phonotactic constraints, phonological alternations occur when morphemes or words come in contact, and depend on the surrounding phonological context. Our stimuli were designed to capture this important aspect. Thus, we asked whether infants could learn that a particular segment surfaces as voiceless when it is preceded by a voiceless consonant, but as voiced in other phonological contexts. We predicted that if infants were able to learn the familiarized alternations, listening preferences for novel sequences at test would differ as a function of the alternation learned.

In addition to determining whether infants are capable of learning alternations from patterns of complementary distribution, we asked when this ability might emerge. The latter half of the first year is a critical period for phonological tuning.

---

[2] The alternations we used are compatible with other characterizations as well (e.g., intervocalic voicing). For convenience, however, we refer to the process as devoicing after voiceless consonants.

For example, sensitivity to native phonotactic patterns emerges around 9 months of age (Jusczyk et al., 1994) and statistical learning of phonotactic patterns has been observed at the same age (Saffran & Thiessen, 2003). Similarly, major changes in consonant perception do not occur before 10 months. Therefore, we might not expect younger infants to succeed at our alternation-learning task. However, infants have shown sensitivity to transitional probabilities in other tasks of artificial sequence learning (e.g., Saffran et al., 1996) and an ability to use distributional information to construct phonetic categories by 6–8 months (Maye et al., 2002). Perhaps the ability to learn from patterns of complementary distribution is in place early as well.

A third possibility is that younger infants will show "partial" learning of these patterns. Statistical learning of phonological alternations may involve two components, requiring different forms of computation. It is possible that the ability to perform these two types of computation emerges in stages: first, noting the relationship between particular sounds and their conditioning contexts, and second, grouping sounds that occur in complementary distribution into a single category. Given their success in tasks of transitional probability learning, infants may be able to note statistical relations between sounds and their conditioning contexts at an early age. However, there are no data concerning infants' ability to categorize elements occurring in complementary distribution; this type of ability may not arise until a later point in development. For this reason, we tested infants of two ages, 12 and 8.5 months.

2. Experiment 1

Two groups of infants were briefly familiarized with artificial languages composed of monosyllabic "determiners" and disyllabic "noun" pairs. Although we refer to the monosyllables and disyllables as determiners and nouns, these labels are arbitrary. Infants were given no evidence as to the lexical status of these nonce syllables. However, these monosyllables and disyllables can be conceptualized more generally as (function word) + (content word) pairs for three reasons. First, the disyllable (noun) set comprised many more items than the monosyllable (determiner) set. Second, disyllables were always preceded by monosyllables. Although this word order is not universal, in English function words precede content words, so this was consistent with the prosodic properties with which our English-learning infants were familiar. Finally, function words tend to be shorter than content words cross-linguistically; they are not subject to word minimality constraints (McCarthy & Prince, 1995; Shi, Morgan, & Allopenna, 1998), consistent with the distinction between our monosyllables and disyllables (e.g., the monosyllable *na* is not a minimal word, as it lacks a coda and contains a short vowel).

For infants in the STOP group, the voicing of initial *stops* in the nouns was conditioned by the voicing of the final segment of the preceding determiner, e.g., *na bevi*, *rot pevi*. In this familiarization language, voiced- and voiceless-initial fricatives occurred freely, regardless of the voicing of the final segment of the preceding deter-

miner. For infants in the FRICATIVE group, the voicing of initial *fricatives* was conditioned by the voicing of the final segment of the determiner, e.g., *na zuma*, *rot suma*. In this familiarization language, voiced- and voiceless-initial stops occurred freely, regardless of the voicing of the final segment of the preceding determiner. Thus half of the infants learned that voiced and voiceless stops were in complementary distribution, whereas the other half learned that voiced and voiceless fricatives alternated.

Following the exposure period, infants heard the same "determiners" *na* and *rot* paired with new "nouns". Nouns occurred in pairs, differing only in the voicing of the initial consonant, which was voiceless following *rot* and voiced following *na*. Half of the nouns began with stop consonants; the other half began with fricatives. All forms were grammatical in both languages. If they had learned the alternations, infants in the STOP group should have interpreted sequences with stop-initial nouns as containing a single, alternating, novel noun, but sequences with fricative-initial nouns as containing two novel nouns. The reverse should have been true for infants in the FRICATIVE group.

2.1. Participants

Sixteen full-term, English-exposed 12-month-olds participated (mean age = 369 days, range 350–388 days). Participants were recruited from Rhode Island birth records and advertisements. Eight infants were assigned to the STOP familiarization group; eight were assigned to the FRICATIVE familiarization group (see below). Data from seven additional infants were not included (five for squirminess and two for significant non-English exposure).

2.2. Stimuli

The artificial languages contained two monosyllables (*na* and *rot*) and 24 disyllables. Sixteen disyllables were presented during familiarization, the remainder during the test phase. Half of the disyllables were stop initial and half began with fricatives. The full set of initial segments was {p, b, t, d, f, v, s, z}. The 24 disyllables consisted of 12 pairs of non-words differing only in the voicing of the initial segment (see Appendix A).

Each disyllable was preceded by one of the two monosyllables. However, the two groups of infants were exposed to different pairings between monosyllables and disyllables during familiarization. For infants in the STOP group, disyllables starting with a voiced stop occurred only after *na*; those starting with a voiceless stop occurred only after *rot*. Infants heard two tokens of each voiced-stop-initial disyllable preceded by *na* and two tokens of each voiceless-stop-initial disyllable preceded by *rot*. By contrast, disyllables starting with fricatives followed both *na* and *rot*, regardless of voicing. Infants heard one token of each voiced-fricative-initial disyllable presented with *na* and one token with *rot*. Similarly, one token of each voiceless-fricative-initial disyllable was presented with *na* and the other with *rot*. Thus, infants in this group had explicit evidence that the voicing alternation did not apply to fric-

atives. Conversely, for infants in the FRICATIVE group, disyllables starting with a stop occurred after both *rot* and *na*, while those starting with a fricative occurred after *rot* if the fricative was voiceless and after *na* if it was voiced. Thus, infants in this group had explicit evidence that the voicing alternation did not apply to stops. By virtue of this design, the frequency of occurrence of all tokens was balanced between the two groups. Infants in both groups heard two tokens of each stop-initial disyllable and two tokens of each fricative-initial disyllable, as well as equivalent numbers of *na* and *rot* tokens. Thus, the only difference between the two groups was the context in which the two sets of voiced- and voiceless-initial consonants occurred. Infants who heard alternations of stop consonants might learn to interpret minimally different stop-initial disyllables as phonological variants of the same noun; however, they should interpret minimally different fricative-initial disyllables as different nouns. The reverse would be true of infants who heard alternations of fricative consonants.

Infants in both groups were also exposed to voiced fricatives and stops intervocalically during familiarization. Specifically, both groups heard two stops ([b], [d]) and two fricatives ([v], [z]) noun-medially. Thus, word-medially infants received no evidence about the alternations – voiced medial segments were chosen so as to conform to both languages. As a result, the alternations could only have been learned by attending to the juncture between the mono- and disyllables.

Stimuli for the test phase consisted of eight monosyllable + disyllable pairs; the monosyllables were the same as during familiarization but the disyllables comprised four pairs of new stop- and fricative-initial non-words differing only in initial voicing. Disyllables beginning with a voiced stop or fricative were preceded by *na*; disyllables beginning with a voiceless stop or fricative were preceded by *rot*. Thus, all monosyllable + disyllable pairings were "grammatical" according to both voicing alternations. Each test sequence involved a single pair of disyllables (e.g., "*rot poli*, *na boli*..."). Two test sequences contained stop-initial disyllables and two contained fricative-initial disyllables. All infants heard the same test sequences. If they had learned the familiarization alternations, each group should interpret one set of sequences as involving a single novel alternating noun; the other set of sequences should be interpreted as involving two novel nouns.

Stimuli were recorded by a trained speaker in an infant-directed register. The stimuli were recorded naturally as phrases consisting of a monosyllable followed by a disyllable. As is the case in natural productions of function + content word sequences, the final [t] of the *rot* determiner was unreleased and there were no pauses between the mono- and disyllables. Cases of natural assimilation typically occur across minor prosodic boundaries, like word boundaries, but not across major phonological or intonational phrase boundaries (Cooper & Paccia-Cooper, 1980; Holst & Nolan, 1995; Nespor & Vogel, 1986; Scott & Cutler, 1984; Selkirk, 1984). Lexical stress was placed on the first syllable of the disyllable. Each disyllable was recorded multiple times with each monosyllable. Two tokens of each monosyllable–disyllable pair were selected for use as stimuli. These stimuli were used for all four experiments.

Despite the lack of pauses between the mono- and disyllables, it was our intent that infants perceive a word boundary between them, allowing us to examine the

learning of phonological alternations, rather than static word-internal phonotactic patterns. Two properties of our stimuli encourage the analysis of these sequences as a monosyllable followed by a separate disyllable. First, there were only two mono-syllables (*rot* and *na*). These two monosyllables could be followed by any of 16 disyl-lables. As a consequence, there was a much lower transitional probability between the monosyllable and the first syllable of the disyllable than between the two syllables of the disyllable. For the stop-alternation familiarization group the transitional probability between the monosyllable and first syllable of the disyllable was .0625 for fricative-initial disyllables and .125 for stop-initial disyllables; the reverse was true for the fricative-alternation familiarization group. The probability between the two syllables of a disyllable was always 1.0. Thus, there was strong statistical evi-dence in favor of a word boundary following the monosyllable (Saffran et al., 1996). Second, the mono- + disyllable sequences were pronounced with wSw stress pat-terns, consistent with the predominant English pattern of a weak function word fol-lowed by a trochaic content word. The acoustic properties of the stimuli – pitch and duration – are consistent with this wSw pattern (see Table 1). Thus, regardless of whether infants were attending to statistical cues, stress cues, or both (Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003), it seemed likely that they would segment these sequences in the way we intended.

2.3. Apparatus

Testing was conducted in a three-walled testing booth within a sound-treated room. Each wall of the booth was 120 cm wide. A chair was positioned at the open end of the booth where parents sat with their infants on their laps. Infants sat approximately 110 cm from the front of the booth. Loudspeakers were located behind both side walls of the booth. At the infants' eye level, a yellow light was mounted on the front wall. Each of the side walls had a similar green light at the same level. A camera was mounted behind the booth just above the center light. A television monitor in a separate control room was connected to the camera in the testing booth. Participants were displayed on the monitor in the control room. The experimenter, who was blind to the experimental condition, watched the mon-itor and recorded the infants' looking by pressing buttons on the mouse of a Win-dows computer to control the customized experimental software. The computer was equipped with a Sound-Blaster compatible soundboard connected to the ampli-fied speakers. The speech was played at a conversational level (75 dB).

Table 1
Stimulus characteristics: mean pitch in Hertz (standard deviation), mean duration in milliseconds (standard deviation) and transitional probabilities

	Syll 1	Syll 2	Syll 3
Pitch	214 (42)	279 (46)	187 (14)
Duration	189 (32)	276 (45)	245 (39)
Transitional probability to next syllable	.0625/.125	1.0	n/a

2.4. Procedure

Infants were tested using a modified version of the Headturn Preference Procedure (Kemler Nelson et al., 1995). Infants sat on their parents' lap facing the center light. Parents listened to instrumental music over headphones to mask the stimuli. Each trial began with the center light flashing until the infant fixated on the light. At that point, this light was turned off, and one of the side lights began to flash to attract the infant's attention to the side. Side of presentation was randomized across trials, so that all stimuli were presented from both sides. After the infant turned to look at the flashing side light, the speech stimuli for that trial began to play.

During familiarization, each trial continued for 40 s after the infant's initial head turn to the side light. Each trial consisted of a randomly ordered presentation of the familiarization stimuli (that is, random presentation of the monosyllable + disyllable pairs); pairs of monosyllables and disyllables were separated by 500 ms intervals (e.g., *rot poli* was followed 500 ms later by *na sadu*). Presentation of stimuli during familiarization was not contingent on the infant's looking behavior. This ensured that looking behavior during the test could not be attributed to differences in the amount of familiarization. There were three familiarization trials, each initiated by the infant's head turn. This made for a total of 120 s of familiarization. Once the sound was initiated, the side light remained on for the duration of each familiarization trial.

The test phase immediately followed the familiarization phase. During this phase, trials were initiated in the same way as in familiarization. However, after initiation of the trial, presentation of the sound was immediately contingent on the infant's looking behavior. Each trial continued until the infant looked away for two seconds, or until 20 s of looking time had been accumulated during that trial. If the infant looked away, but then looked back within two seconds, the trial continued. If the infant's looking time was less than two s, the trial was repeated with a new randomization of the trial stimuli (again, random presentation of the monosyllable + disyllable pairs); otherwise, the procedure advanced to the next trial. During test trials, the side light continued to flash while the speech stimuli were presented, to maintain the infant's interest.

The test phase comprised three blocks of four trials each. Each of the four test sequences was presented once per block. The order of the sequences within blocks was randomized anew for each infant. In addition, the order of the stimuli within sequences was randomized on each trial.

2.5. Results and discussion

If they had acquired the familiarization alternations, infants in the STOP group should have interpreted the test sequences with stop-initial disyllables as involving a single novel noun (the initial consonant of which changed depending on the preceding determiner), but sequences with fricative-initial disyllables as involving two novel nouns. The reverse should have been true of the infants familiarized with the fricative alternations. We expected that, if infants had different interpretations of the test sequences, looking times for the sequences would differ. Because infants' looking

preferences are governed by a host of factors (Hunter & Ames, 1988), we could not predict the direction of preference in advance. However, Maye et al. (2002), in their study of phonetic category learning, found that infants preferred to listen to repetitions of sounds from a single category over alternations of sounds originating from two different categories (which the authors interpreted as a type of novelty preference because infants had heard many different sounds in familiarization). If our study similarly addresses the formation of single vs. multiple sound categories, we might expect a preference for sequences interpreted as containing a single novel noun over sequences containing two novel nouns.

In fact, infants in the STOP group listened longer to sequences with stop-initial disyllables (8.2 s vs. 7.2 s), whereas infants in the FRICATIVE group listened longer to sequences with fricative-initial disyllables (9.4 s vs. 7.4 s). A 2×2 mixed ANOVA with familiarization group (stop or fricative) as a between-subjects factor and test type (stop-initial or fricative-initial) as a within-subjects factor revealed no significant effect of test type, $F(1,14) = .97$, ns, and no effect of familiarization group, $F(1,14) = .37$, ns, but a significant interaction, $F(1,14) = 7.42$, $p < .025$. Each group had a preference for the sequences exemplifying one novel alternating noun as determined by their familiarization stimuli. Fig. 1 displays mean overall looking times for sequences containing one novel noun versus two novel nouns. Overall, 12-month-olds had a significant preference for the sequences exemplifying a single alternating noun (8.8 s vs. 7.3 s), $t(15) = 2.73$, $p < .025$ (two-tailed), Cohen's $d = .68$ (Cohen, 1988). Eleven of 16 infants preferred the sequences that corresponded to one novel noun as determined by their familiarization stimuli.

During familiarization, infants in the two groups heard equivalent numbers of stop- and fricative-initial disyllables and an equivalent number of syllables beginning with voiced and voiceless segments. Therefore, infants' differential preference for the two types of test sequences is consistent with the idea that they functionally grouped alternating voiced and voiceless segments.

Given the developmental trajectory of phonological learning that we outlined at the outset, it is unclear when the ability to detect the patterns of complementary dis-
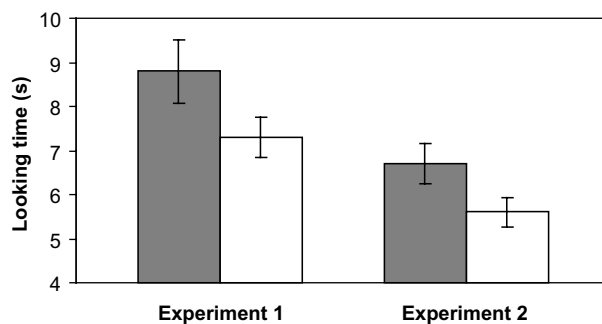

Fig. 1. Summary of results for Experiments 1 (12-month-olds) and 2 (8.5-month-olds): looking times (in seconds) for sequences consisting of one novel (gray bars) or two novel (white bars) nouns.

tribution underlying phonological alternations might emerge. By 12 months of age, infants also show advances in other areas of phonological perception. However, younger infants are impressive statistical learners. This raises the question of whether younger infants can also learn novel alternations. To address this question, we tested 8.5-month-olds on the same task.

3. Experiment 2

In Experiment 2, 8.5-month-olds were tested using the stimuli and procedure of Experiment 1 to explore whether the ability to quickly learn novel alternations is present at an earlier stage of development. Given increases in processing speed over the first two years (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998) and evidence that slowed speech enhances young infants' performance (Morgan et al., 2002), we hypothesized that longer intervals between monosyllable–disyllable pairs would help younger infants track the relationship between alternating segments.[3] We thus increased the inter-stimulus pair interval in Experiment 2.

3.1. Participants

Twenty-six full-term, English-exposed 8.5-month-olds participated (mean age = 277 days, range 257–290 days). Thirteen infants were assigned to the STOP familiarization group; thirteen were assigned to the FRICATIVE group. Data from 10 additional infants were not included (eight for squirminess/fussiness, one for disinterest, and one for non-English exposure).

3.2. Stimuli, apparatus and procedure

The stimuli and procedure were identical to Experiment 1 (the same recordings were used), with the following exceptions: the interval *between* monosyllable–disyllable pairs was increased from 500 ms to 1 s in both familiarization and test (e.g, *rot poli* was followed one second later by *na sadu*). The transitions between mono- and disyllables within pairs were not altered. To ensure that infants heard the familiarization stimuli for a sufficient amount of time, there was one additional 40-s familiarization trial, for a total of 160 s. Infants in both Experiment 1 and Experiment 2 heard approximately 100 pairs of monosyllables + disyllables during the familiarization phase.

---

[3] We first conducted this study using the same inter-pair interval as in Experiment 1 (500 ms between monosyllable + disyllable pairs). Twenty-six English-exposed 8.5-month-olds were tested using the stimuli and procedure of Experiment 1 (mean age = 260 days, range 245–278 days). Infants exhibited a non-significant preference for sequences exemplifying a single novel alternating noun (7.9 s vs. 7.2 s).

3.3. Results and discussion

Infants' looking times are shown in Fig. 1. Infants in the STOP group listened longer to the stop-initial test sequences (7.3 s vs. 6.1 s). Infants in the FRICATIVE group, conversely, listened longer to the fricative-initial test sequences (6.1 s vs. 5.2 s). These differences were in the same direction as those observed for 12-month-olds in Experiment 1. A 2×2 mixed ANOVA with familiarization group (stop or fricative) as a between-subjects factor and test type (stop-initial or fricative-initial) as a within-subjects factor found no significant effect of test type, $F(1, 24) = .16$, *ns*, or familiarization group, $F(1, 24) = 2.46$, $p > .1$, but a significant interaction, $F(1, 24) = 5.29$, $p < .05$. Each group had a preference for the sequences exemplifying one novel alternating noun as determined by their familiarization stimuli. Overall 8.5-month-olds had a significant preference for sequences exemplifying a single alternating noun (mean 6.7 s vs. 5.6 s), $t(25) = 2.34$, $p < .05$ (two-tailed), Cohen's $d = .46$. Seventeen out of 26 infants showed this pattern. The results of this study suggest that younger infants are also able to learn novel phonological alternations, perhaps even functionally grouping alternating segments.

However, as pointed out in the introduction, learning phonological alternations from distributional cues might involve two stages, the first of which is noting the statistical relationship between particular sounds and their conditioning contexts. Infants might have exhibited this pattern of performance not because they grouped the alternating sounds into a single functional category, but because they detected dependencies between sounds and their conditioning contexts. For each group of infants, test sequences constituting a single novel noun also contained determiner + noun onset sequences of higher transitional probability than test sequences constituting two novel nouns.

To illustrate this point, consider infants in the STOP group. For these infants, voiced stops were always preceded by the [a]-final determiner, whereas voiceless stops were always preceded by the [t]-final determiner. However, voiced fricatives occurred with both [a]-final and [t]-final determiners (as did voiceless fricatives). Therefore, during familiarization, the transitional probability of voiced stops given an [a]-final determiner and voiceless stops given a [t]-final determiner was twice the transitional probability of voiced fricatives given an [a]-final determiner and voiceless fricatives given a [t]-final determiner. At test, all sequences were grammatical for both groups; in other words, all voiced obstruents were preceded by the [a]-final determiner and all voiceless obstruents were preceded by the [t]-final determiner. Hence, the preference infants in the STOP group expressed for stop-initial sequences at test could have been due to the higher probability with which those sequences occurred during familiarization.

If infants performed the task by attending to the probabilities with which different transitions occurred, this raises the issue of what sized units might be involved in these computations. Most previous demonstrations that infants are sensitive to transitional probabilities have involved cases where both syllabic and segmental transitions predicted the same outcome (e.g., Saffran et al., 1996). Indeed, although even newborns can detect syllable-sized units (Bijeljac-Babic, Bertoncini, & Mehler,

1993), it is not clear at what point infants begin to represent segments. Jusczyk and Derrah (1987) explored the nature of 2-month-olds' representations for speech sounds by habituating different groups of infants to a set of syllables beginning with a particular consonant, [b]. At test, some infants were presented with CV syllables differing in the vowel alone ([bu]) and other infants were presented with CV syllables differing in both the consonantal and vocalic portions ([du]). Jusczyk and Derrah reasoned that if infants' representations contain information about individual segments, the syllables differing in both the consonant and vowel should be perceived as more novel and lead to greater dishabituation (analogous to between-category vs. within-category changes in tests of visual perception). Results showed that dishabituation to CV syllables differing in the consonant alone was of the same magnitude as dishabituation to syllables differing in both the consonantal and vocalic portions. Thus, Jusczyk and Derrah concluded that infants may not initially analyze speech in terms of segments. Similarly, Eimas (1999) familiarized 3- to 4-month-old infants with sets of syllables beginning with either [b] or [d] and then tested them on their preference for novel syllables beginning with one sound or the other. Finding that infants did not prefer test items beginning with the familiar segment to those beginning with the novel segment, Eimas concluded that infants at this age do not yet categorize segments.

In our stimuli, with one pair of exceptions, test nouns did not share syllables with familiarization nouns. These exceptions are *poli* and *boli*, which overlapped with the familiarization nouns *pogu* and *bogu*. To test whether the observed preference for single alternating nouns might have been due to these particular stimuli alone, we excluded them, comparing listening times in the STOP group for the *pazo-bazo* stimuli (8.3 s) to listening times for the fricative-initial sequences (6.1 s). This difference was significant, $t(12) = 3.54$, $p < .01$, suggesting that by 8.5 months, infants are sensitive to patterns of segments, as well as syllables. Relatedly Newport, Weiss, Wonnacott, and Aslin (2004) have shown that 8-month-old infants can segment continuous streams of speech when segmental and syllabic transitional probabilities both have minima at 'word' boundaries, but not when these two types of minima occur in differing locations – suggesting that 8-month-old infants are tracking transitions at both levels.

The results of Experiments 1 and 2 demonstrate that, at a minimum, infants learned the probability with which different phoneme sequences occurred in the familiarization stimuli. However, they do not indicate whether the infants went beyond transitional probabilities to group alternating segments. To determine whether infants' performance could be completely explained by attention to transitional probabilities, we familiarized infants as before, but tested them on novel bare nouns: determiners were omitted from test stimuli, removing any transitional probabilities that might distinguish the two test sets.

4. Experiment 3

In this experiment, we sought evidence that 8.5-month-old infants can not only learn about the relative probabilities of different determiner + noun onset sequences,

but can also group alternating segments into a single category. In Experiments 1 and 2, as in Maye et al. (2002), infants preferred the single-item test stimuli. We expected them to do the same here and adjusted our statistical tests accordingly.

4.1. Participants

Twenty-six full-term, English-exposed 8.5-month-olds participated (mean age = 266 days, range 243–295 days). Thirteen infants were assigned to the STOP familiarization group; thirteen were assigned to the FRICATIVE group. Data from five additional infants were not included (four for fussiness, one for disinterest in the lights).

4.2. Stimuli, apparatus and procedure

The procedure and familiarization stimuli were identical to Experiment 2. Again, the interval between pairs of mono- and disyllables was one second. The difference arose in the test phase: unlike the previous experiments, here there were no determiners at test. This was accomplished by removing the determiners from the stimulus recordings used in Experiment 2. The portion of the waveform prior to the disyllable's onset (stop burst or onset of frication) was excised. Removing the determiner did not introduce any clicks or other artifacts into the signal. The interval between the disyllables during test was 1 s. We removed the determiners so that infants' test preferences could not be driven by the relative probability of local determiner-onset transitions in familiarization. This, therefore, constitutes a more stringent test of our hypothesis that infants can group alternating segments into a single category. Even if infants do track the transitional probabilities of alternating segments and their conditioning contexts during familiarization, this would not allow them to differentiate the test sequences when determiners are missing. Therefore, continued differentiation of the test sequences would have to be driven by additional information extracted during familiarization, specifically, information about the patterns of complementary distribution.

4.3. Results and discussion

Infants' looking times are shown in Fig. 2. Infants in the STOP group displayed no preference for either test sequence (6.7 s for stop-initial sequences vs. 6.6 s for fricative-initial test sequences); infants in the FRICATIVE group had slightly higher looking times for fricative-initial sequences (6.9 s vs. 6.4 s). A 2×2 mixed ANOVA with familiarization group (stop or fricative) as a between-subjects factor and test type (stop-initial or fricative-initial) as a within-subjects factor revealed no significant effects (test type, $F(1, 24) = .24$, *ns*; familiarization group, $F(1, 24) = .002$, *ns*; interaction $F(1, 24) = .43$, *ns*). Overall there was a small, non-significant preference for sequences exemplifying a single alternating noun (6.8 s vs. 6.5 s), $t(25) = .67$, *ns*, Cohen's $d = .13$. Thirteen of 26 infants exhibited a preference for single-noun sequences.
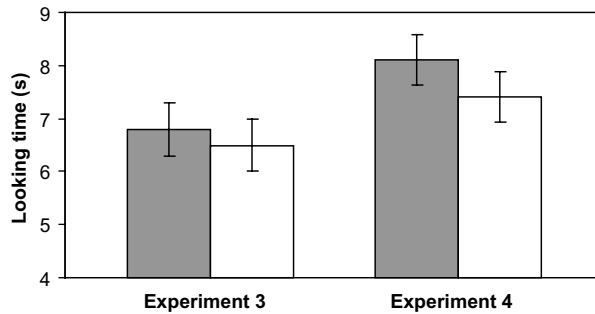
Fig. 2. Summary of results for Experiments 3 (8.5-month-olds) and 4 (12-month-olds): looking times (in seconds) for sequences consisting of one novel (gray bars) or two novel (white bars) nouns.

To compare 8.5-month-olds' performance in Experiments 2 and 3, we conducted a $2 \times 2 \times 2$ mixed ANOVA with test sequence (one novel or two novel) as a within-subjects factor and familiarization group (stop or fricative) and Experiment (2 or 3) as between-subject factors. There was a main effect of test sequence ($F(1,48) = 4.74$, $p < .05$). No other effects or interactions were significant (experiment $F(1,48) = .74$, ns; familiarization group $F(1,48) = .81$, ns; experiment × familiarization group $F(1,48) = .69$, ns; experiment × test sequence $F(1,48) = 1.75$, $p < .19$; familiarization group × test sequence $F(1,48) = 0$, ns; experiment × familiarization group × test sequence $F(1,48) = .38$, ns). Although the interaction of experiment and test sequence did not reach significance, 8.5-month-olds' performance across the two experiments suggests that they are able to learn about the statistical relationship between sounds and their conditioning contexts, but are not yet able to group segments which occur in different conditioning contexts.

Younger infants failed to express a preference when transitional probabilities were unavailable at test. However, it is possible that older infants are able to extract information in addition to transitional probabilities given their greater experience and phonological sophistication. Therefore, in Experiment 4 we examined whether 12-month-olds can go beyond this type of statistical dependency to group alternating segments into a single functional category.

5. Experiment 4

5.1. Participants

Twenty-four full-term, English-exposed 12-month-olds participated (mean age = 365 days, range 342–384 days). Twelve infants were assigned to the STOP familiarization group; twelve were assigned to the FRICATIVE group. Data from four additional infants were not included (three for squirminess/fussiness and one because the infant was 6 weeks premature).

5.2. Stimuli, apparatus and procedure

Identical to Experiment 1, but with the determiners removed from the test sequences as in Experiment 3. As in Experiment 1, the interval between pairs of mono- and disyllables was 500 ms in familiarization. The interval between isolated disyllables was 500 ms in test.

5.3. Results and discussion

Infants' looking times are shown in Fig. 2. Infants in the STOP group listened longer to the stop-initial test sequences (7.2 s vs. 6.6 s). Infants in the FRICATIVE group, conversely, listened longer to the fricative-initial test sequences (9.0 s vs. 8.1 s). A 2×2 mixed ANOVA with familiarization group (stop or fricative) as a between-subjects factor and test type (stop-initial or fricative initial) as a within-subjects factor revealed a marginally significant effect of familiarization group ($F(1,22) = 3.96$, $p < .06$; overall, infants in the FRICATIVE group listened to the test stimuli for 8.6 s vs. 6.9 s for infants in the STOP group), no effect of test type ($F(1,22) = .27$, *ns*), and a marginally significant interaction ($F(1,22) = 3.31$, $p < .08$). Overall this led to a significant preference for sequences exemplifying a single alternating noun (8.1 s vs. 7.3 s), $t(23) = 1.85$, $p < .05$ (one-tailed), Cohen's $d = .38$. Sixteen out of 24 infants had longer listening times for sequences exemplifying one novel noun.

To assess whether 12-month-olds' performance differed in Experiments 1 and 4, we conducted a $2 \times 2 \times 2$ mixed ANOVA with test sequence (one novel or two novel) as a within-subjects factor and familiarization group (stop or fricative) and Experiment (1 or 4) as between-subject factors. There was a significant effect of test sequence ($F(1,36) = 11.25$, $p < .002$) and a marginal effect of familiarization group ($F(1,36) = 2.93$, $p < .1$). No other effects or interactions were significant (experiment $F(1,36) = .2$, *ns*; experiment $\times$ familiarization group $F(1,36) = .53$, *ns*; experiment $\times$ test sequence $F(1,36) = 1.36$, $p < .25$; familiarization group $\times$ test sequence $F(1,36) = 1.28$, $p < .27$; experiment $\times$ familiarization group $\times$ test sequence $F(1,36) = .25$, *ns*). Twelve-month-olds' successful discrimination in both Experiments 1 and 4 suggests that they can go beyond transitional probabilities to start forming functional groupings of alternating segments.

To assess further whether 12-month-olds' performance in Experiment 4 differed from 8.5-month-olds' performance in Experiment 3, we conducted a $2 \times 2 \times 2$ mixed ANOVA with test sequence (one novel or two novel) as a within-subjects factor and familiarization group (stop or fricative) and age (8.5 or 12 months) as between-subject factors. There were marginally significant effects of test sequence ($F(1,46) = 3.1$, $p < .09$) and of age ($F(1,46) = 2.88$, $p < .1$). No other main effects or interactions were significant (familiarization group $F(1,46) = 1.59$, $p < .21$; age $\times$ familiarization group $F(1,46) = 1.74$, $p < .19$; age $\times$ test sequence $F(1,46) = .71$, *ns*; familiarization group $\times$ test sequence $F(1,46) = .51$, *ns*; age $\times$ familiarization group $\times$ test sequence $F(1,46) = .001$, *ns*). The results of the individual experiments suggest that infants first become sensitive to the relationship between sounds and their conditioning contexts, and later develop the ability to group similar sounds occurring in complemen-

tary conditioning contexts. However, because the interaction of age and test sequence was not significant, this conclusion must be drawn cautiously.

6. General discussion

In the present studies, we explored one strategy by which young learners might learn about the phonological alternations of their native language. Infants were exposed to artificial languages in which the only information that signaled the presence of phonological alternations was distributional. Infants were exposed to either stop or fricative voicing alternations. Following this brief exposure period, their preferences for novel sequences involving voiced and voiceless stops and fricatives were assessed. Both 12- and 8.5-month-olds exhibited differential treatment of alternating and non-alternating segments in their familiarization language. However, younger infants discriminated alternating and non-alternating test sequences only when the conditioning contexts (i.e., the monosyllabic 'determiners') were present. The 12-month-olds, in contrast, exhibited the same pattern of performance regardless of whether or not the conditioning contexts were present.

In the three experiments in which infants expressed a significant preference, it was for sequences containing what could be conceptualized as a single novel noun appearing in alternating forms over sequences containing two novel nouns. This pattern of performance is consistent with the direction of preference exhibited by infants in Maye et al. (2002). In that study, infants were familiarized with sounds from a [da]-[ta] continuum presented in frequency distributions characteristic of either a single (unimodal group) or multiple (bimodal group) phonetic categories. At test, all infants heard two types of trials: trials containing a single repeating sound and trials in which two sounds alternated. Only infants in the bimodal familiarization group should have interpreted the latter type of test trial as involving sounds from two different phonetic categories. For infants in the unimodal group, these same sounds should have been interpreted as originating from the same phonetic category. Results showed that infants in the bimodal familiarization group expressed a preference for a single repeating sound over sounds that should have been interpreted as belonging to two separate categories, but those in the unimodal group had no preference. In the absence of training, infants presumably would have discriminated the sounds and therefore shown a preference for one of the test sequences; Maye et al. (p. B104) note that their stimulus continuum was "based on a phonetic contrast that infants between 6 and 8 months of age have been shown to discriminate: voiced unaspirated vs. voiceless unaspirated stop consonants (Pegg & Werker, 1997)". Therefore, the effect of the training was likely in the unimodal condition: hearing two sounds drawn from a single distribution caused infants to collapse these sounds into the same category; as a result, infants in this group no longer discriminated the sounds.

In our studies, we used a counterbalanced design in which, during the test phase, each infant heard two stimulus pairs from one underlying category (e.g., nouns beginning with voiced and voiceless stops) and two stimulus pairs from two underlying categories (e.g., nouns beginning with voiced and voiceless fricatives). Similar to Maye et al., we found a preference for sequences in which application of the

familiarization alternation should have caused infants to interpret test stimulus pairs as belonging to the same category. The parallels between these two sets of results suggest that our infants, like those in the Maye et al. study, categorized sounds differently as a function of the familiarization patterns. Moreover, they are consistent with the possibility that older infants in our study perceived alternating sounds as originating from a single category. In other words, as in Maye et al., training may have led infants to collapse discriminable sounds into one category. A crucial difference between the two studies, however, is that in Maye et al., infants' preference was for a single surface sound; in our studies, infants heard the same number of surface sounds in the test sequences but preferred listening to sounds from a single *underlying* category. Thus, if our interpretation is correct, this type of task can be extended to investigate questions of *functional* grouping, and not simply grouping based on surface similarity. As a result, this task offers a useful tool for exploring the nature of infants' underlying representations for speech sounds.

A preference for sounds originating from a single category would emerge only if 12-month-olds learned the familiarization alternations. However, acquisition of the familiarization alternations could also produce the pattern of results in Experiment 4 through a different route. Rather than displaying a familiarity preference for a single category during test, 12-month-olds might have been displaying a type of novelty preference due to perceived ungrammaticality in certain test sequences. To see this, consider infants in the STOP group, who were familiarized with a language in which word-initial stops were voiceless following a voiceless consonant and voiced elsewhere. For infants in this group, a test sequence like *boli poli boli boli poli*... contains an illegality, namely, that voiceless-initial *poli* is preceded by a vowel (that is, the vowel at the end of the preceding word *boli* or *poli*); the same type of illegality would not be present in the fricative test sequences. Thus, rather than showing a preference for "one word" test sequences, infants may have been listening longer to sequences containing the violation. This preference for sequences violating the familiar pattern would constitute a type of novelty preference. Such a flip in preference (if infants were exhibiting a preference for highly probable transitions in Experiments 1 and 2) could be explained by the greater simplicity of the test stimuli in Experiments 3 and 4 (Hunter & Ames, 1988). However, an account of infants' preferences that relies on this detection of contextual ungrammaticality seems unlikely given the long interval between adjacent disyllables in test (500 ms). As mentioned earlier, alternations of this sort typically occur within phonological phrases (as in the familiarization stimuli used here), not across phonological phrase boundaries. Nevertheless, if infants in the STOP group did in fact learn that *boli* was the underlying form, and that *poli* surfaced only when preceded by a voiceless consonant within the same phonological phrase, then the occurrence of *poli* after a 500-ms pause (that is, in the absence of any conditioning environment) could also be considered a violation.[4]

---

[4] Note that we have chosen to characterize the alternation as devoicing following a voiceless consonant; in this argument, we assume that the underlying form is *boli* and the derived form is *poli*. If infants internalized a different alternation (e.g., intervocalic voicing) the issue of illegality in the test sequences would remain, mutatis mutandis.

Of course, this account requires that infants have learned the illustrated alternations and that they have collapsed the two surface manifestations into a single underlying category.

We cannot at present determine which of these explanations is correct – whether 12-month-olds' success in Experiment 4 is explained by a change in their perception of alternating segments or by detection of a violation. Again, however, it is important to note that both of these explanations rely on infants having learned the alternations during the familiarization phase. At the same time, they may have different implications for how infants represent and perceive alternating segments. Whereas both accounts suggest that exposure to this type of distributional information leads infants to group alternating segments into the same functional category, only the first account suggests further that infants may perceive alternating segments as more similar (as is the case in adults' allophonic perception; Kazanina et al., 2006; Whalen et al., 1997).

The present results suggest that distributional learning is a viable strategy for acquiring phonological alternations, like allophony or allomorphy, that are characterized by patterns of complementary distribution. To our knowledge, this is the first demonstration that infants are capable of learning patterns of complementary distribution, and is consistent with the growing body of work demonstrating that infants are highly capable statistical learners (Gómez, 2002; Maye et al., 2002; Saffran et al., 1996).

In addition, these results suggest a developmental trajectory for this type of learning. We hypothesized that learning phonological alternations from distributional cues involves two components: noting the statistical relationship between sounds and their conditioning contexts and then grouping sounds that occur in different conditioning contexts. The results of the present studies suggest that only 12-month-olds form functional groupings of alternating sounds, as they successfully differentiated the test sequences whether or not information about transitional probabilities was available at test. Thus, 1-year-olds can learn novel phonological alternations from as little as two minutes of exposure on the basis of distributional information alone. In contrast, 8.5-months-olds appear to have detected only the relationship between sounds and their conditioning contexts. If indeed 8.5-month-olds were reliant on transitional probabilities (or on the overt presence of conditioning contexts as cues), this is consistent with other demonstrations that infants at this age can track the relationship between syllables and segments (Newport et al., 2004; Saffran et al., 1996), and other, non-linguistic stimuli (Kirkham, Slemmer, & Johnson, 2002). It is important to note that if the younger infants were relying on transitional probabilities, it was transitions between segments rather than, or perhaps in addition to, syllables. Thus these results and future work along these lines can contribute to our understanding of when in development infants represent speech at various levels of detail (e.g., segment, syllable) and over which units infants perform computations on speech input.

Moreover, given that infants' perception of native and non-native phonetic categories changes in the latter half of the first year (Anderson et al., 2003; Werker & Tees, 1984), these results suggest that the acquisition of alternations may occur

in parallel with other aspects of phonological acquisition. In other words, infants are not only learning about phonetic categories and static phonotactic patterns during the first year, but may also be beginning to group phones into phonemic categories and possibly learning other types of phonological alternations as well. Complicating the task for the learner, of course, is that two segments that alternate at one level (e.g., allomorphs or segments related by sandhi processes) may be contrastive at another level (phonemes). Thus the learner must be sensitive to and coordinate multiple sets of distributions. Additional processes, such as morpheme segmentation, may be critical for determining at what level an alternation occurs.

What sort of alternation did infants learn in our experiments? In English, the complementary distributions that define allophonic alternations typically involve differences with respect to syllable position or syllabic stress. For example, in English, the aspirated allophone $[p^h]$ occurs syllable initially, whereas the unaspirated allophone $[p]$ occurs in syllable onsets following $[s]$. Our stimuli did not incorporate such variations in syllable position. However, allophonic alternations are not necessarily conditioned by syllabic position in other languages. Therefore, given the simple structure of our stimuli, the alternation that infants learned could be characterized as allophonic, allomorphic or allolexic, depending in part on whether infants perceived a word boundary between the determiner and noun.

It is likely that alternations differ in the ease with which they can be learned. Some alternations involve segments that must be assigned to functionally distinct categories at one level, but considered functionally equivalent at another level. For example, morphophonological and sandhi alternations involve segments that are functionally distinct at the phonemic level. As a result, listeners cannot simply collapse these segments indiscriminately across the board. Instead, they must track statistics and function across multiple levels simultaneously. And even in the case of allophonic alternations, learners must maintain some sensitivity to allophonic variants, as these are important cues to word segmentation (Jusczyk et al., 1999). Alternations also differ with respect to what information is relevant: for example, as noted earlier, information about phonological phrase boundaries is important for determining whether a sandhi alternation is licensed. In other cases, information about syllable position, syllable boundaries or morpheme boundaries is critical.

Learning of alternations may benefit other aspects of acquisition. Several theorists have suggested that learners operate under a principle of avoiding exceptions (MacWhinney, 1978; Slobin, 1973, 1985). MacWhinney (1983) noted,

> …this means that a child will attempt to acquire a single form to express a single function. Furthermore, once one form has been learned, the child should resist acquisition of a second, synonymous form. In other words, *when the alternation between two forms cannot be predicted* by any combination of phonological or semantic conditions, acquisition of the second form of the pair should be slow and prone to error. (p. 468, emphasis ours).

The obvious corollary of MacWhinney's assertion is that understanding alternations will allow learners to reduce multiple forms to a single form and thus map this single form to a single function.

For example, knowing the allolexic pattern according to which word-final nasals assimilate in place to following word-initial obstruents can help the learner to avoid positing a spurious new lexical entry "browm" upon hearing the sequence "browm bear". In addition, knowing the conditions under which allomorphs alternate can help to avoid morphological oversegmentation. Whereas the sequences [roz] and [laks] are ambiguous, analyzable as either mono-morphemic or bi-morphemic forms, the sequence [rIns] is not, because after [n] the plural allomorph is [z]. Once the distribution of plural allomorphs is understood, [rIns] can only be analyzed as mono-morphemic. Inclusion of such knowledge of phonological alternations may improve the performance of inferential models of morpheme discovery (Goldwater, Griffiths, & Johnson, 2006).

One question that remains is whether the type of statistical learning observed here is constrained by phonological principles, simplifying the learning problem for infants. There are at least two ways in which these types of alternations are phonologically constrained. First, many phonological processes are motivated by articulatory factors and are thus phonetically natural[5]: these alternations involve closely related sounds in contexts that phonetically motivate the alternations (e.g., the [z] plural allomorph surfaces as the closely related voiceless [s] allomorph when adjacent to a voiceless segment). Simulation work on the acquisition of allophonic rules suggests that a filtering mechanism based on phonetic naturalness may be necessary to prevent the acquisition of spurious alternations (Peperkamp et al., 2006a). Second, phonological alternations often involve several segments that form a natural class and likewise apply within contexts that form a natural class. That is, related sounds (e.g., voiced stops) often undergo the same processes and these processes are often triggered by sounds that are themselves related.

Previous work on the role of these factors in adults' acquisition of new alternations has led to mixed results. For example, with respect to phonetic naturalness, adults' ability to learn unnatural processes appears to differ as a function of the type (or degree) of unnaturalness and task (Peperkamp, Skoruppa, & Dupoux, 2006b). With respect to the issue of natural classes, generalization within a natural class was not observed in Peperkamp and Dupoux (2007) or in Peperkamp et al. (2006b). In Wilson (2006), however, an asymmetric generalization pattern was found for triggering segments; generalization occurred only from a less phonetically natural trigger to a stronger one. Specifically, velar palatalization, a process by which velar consonants become fronted, was generalized to an [i] environment when learned in an [e] environment; [i] is phonetically more front than [e] and hence a more natural context for palatalization; generalization was not observed in the other direction (i.e., from [i] to [e]). Thus phonetic naturalness may affect the degree to which alternations are generalized.

---

[5] Other phonological processes (e.g. dissimilation) may be motivated by perceptual factors.

Thus far, no research has considered the potential roles of phonetic naturalness and natural classes in infants' acquisition of phonological alternations. However, there has been some work on the role of these factors in related areas of phonological acquisition. For example, in the case of phonetic category formation, infants, as opposed to adults, have been observed to generalize to other members of a natural class (Maye & Gerken, 2001; Maye & Weiss, 2003). Also, Saffran and Thiessen (2003) demonstrated that infants may learn phonotactic patterns involving natural classes of segments more easily than patterns involving unrelated groups of segments. In contrast, Seidl and Buckley (2005) found no evidence that the acquisition of phonotactic patterns involving less phonetically natural pairings of sounds and contexts was more difficult than the acquisition of more phonetically natural patterns.

With respect to the present studies, considerations of phonological naturalness raise two additional interpretative questions. First, would infants learn an alternation involving unrelated segments, or conditioning contexts that did not motivate the change? Second, precisely what alternations did infants in our studies learn? Did infants in the STOP groups learn an alternation involving the voiced labial and coronal stops [b] and [d] or involving the natural class of voiced stops? In the latter case, it should be possible to observe generalization of the alternation to the third member of this natural class, i.e., [g]. Similarly, one can ask what infants learned about the conditioning context of the alternation. Manipulations of the nature of alternating segments and the relationship between alternating segments and conditioning contexts can provide further insight into the size of infants' computational units and the role of phonological knowledge.

Infants make considerable progress towards acquiring the sound system, and in particular, the segmental inventory of their native language, during the first year of life. The current results reveal that, in addition, infants have at their disposal a powerful tool for discovering relationships among these segments. Future work will clarify whether infants' structural analyses of the input are unconstrained, or whether infants are instead predisposed to search for patterns that are most likely to facilitate their acquisition of phonological structure, patterns that are attested across human languages.

Acknowledgements

Appendix A

A.1. Familiarization stimuli

Monosyllables: *na*, *rot*
Disyllables:
   Stop-initial: *bevi*, *pevi*, *bogu*, *pogu*, *dula*, *tula*, *dizu*, *tizu*
   Fricative-initial: *zuma*, *suma*, *zobi*, *sobi*, *veda*, *feda*, *vanu*, *fanu*

	STOP group	FRICATIVE group
na	*bevi*, *bogu*, *dula*, *dizu*	*zuma*, *zobi*, *veda*, *vanu*
rot	*pevi*, *pogu*, *tula*, *tizu*	*suma*, *sobi*, *feda*, *fanu*
na, *rot*	*zuma*, *zobi*, *veda*, *vanu* *suma*, *sobi*, *feda*, *fanu*	*bevi*, *bogu*, *dula*, *dizu* *pevi*, *pogu*, *tula*, *tizu*

A.2. Test sequences

Stop-initial
   (1) *rot poli na boli rot poli na boli…*
   (2) *rot pazo na bazo na bazo rot pazo…*
Fricative-initial
   (3) *rot sadu rot sadu na zadu rot sadu…*
   (4) *rot seenay na zeenay na zeenay rot seenay…*

References

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*(2-3), 155–182.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances?. *Developmental Psychology 29*(4), 711–721.

Bortfeld, H., & Morgan, J.L. (submitted for publication). Early word recognition may be stress-full.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science, 16*(4), 298–304.

Chambers, K., Onishi, K., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition, 87*, B69–B77.

Clements, N. (1978). Tone and syntax in Ewe. In D. J. Napoli (Ed.), *Elements of tone, stress, and intonation* (pp. 21–99). Washington DC: Georgetown University Press.

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes, 16*(5–6), 535–564.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, W. E., & Paccia-Cooper, J. M. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.

Darcy, I., Ramus, F., Christophe, A., Kinzler, K., & Dupoux, E. (in press). Phonological knowledge in compensation for native and non-native assimilation. In: F. Kügler, C. Féry & R. van de Vijver (Eds.), *Variation and gradience in phonetics and phonology*. Berlin: Mouton De Gruyter.

Eimas, P. D. (1999). Segmental and syllabic representations in the perception of speech by young infants. *Journal of the Acoustical Society of America, 105*(3), 1901–1911.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*(3), 228–231.

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*(1), 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*(2), 380–396.

Goldwater, S., Griffiths, T. L., & Johnson, M. (2006). Interpolating between types and tokens by estimating power-law generators. *Advances in Neural Information Processing Systems, 18*, 459–466.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*(5), 431–436.

Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception & Psychophysics, 65*(4), 575–590.

Holst, T., & Nolan, F. (1995). The influence of syntactic structure on [s] to [S] assimilation. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in laboratory phonology IV* (pp. 315–333). Cambridge, England: Cambridge University Press.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1570–1582.

Hunter, M., & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.). *Advances in infancy research* (Vol. 5, pp. 69–95). Norwood, NJ: Ablex.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*, 548–567.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology, 29*(1), 1–23.

Jusczyk, P. W., & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology, 23*(5), 648–654.

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.

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics, 61*(8), 1465–1476.

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America, 103*(30), 11381–11386.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development, 18*(1), 111–116.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence of a domain general learning mechanism. *Cognition, 83*(2), B35–B42.

Kouider, S., Halberda, J., Wood, J., & Carey, S. (2006). Acquisition of English number marking: The singular–plural distinction. *Language Learning and Development, 2*(1), 1–25.

Kuhl, P. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development, 6*, 263–285.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606–608.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*(3), 245–294.

MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development, 43* (Vols. 1–2, Serial No. 174).

MacWhinney, B. (1983). Miniature linguistic systems as tests of the use of universal operating principles in second-language learning by children and adults. *Journal of Psycholinguistic Research, 12*(5), 467–478.

Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 644–655.

Maye, J. & Gerken, L. (2001). Learning phonemes: How far can the input take us? In: A. H.-J. Do, L. Dominguez & A. Johansen (Eds.), *Proceedings of the 25th annual Boston University conference on language development* (pp. 480–490). Sommerville, MA: Cascadilla Press.

Maye, J., & Weiss, D. J. (2003). Statistical cues facilitate infants' discrimination 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). Somerville, MA: Cascadilla Press.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101–B111.

McCarthy, J., & Prince, A. (1995). Prosodic morphology. In J. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 318–366). Oxford: Blackwell.

Morgan, J., Singh, L., Bortfeld, H. Rathbun, K., White, K., & Anderson, J. (2002). Infant word recognition: Sentence position and processing time. Paper presented at the *International Conference on Infants Studies*, Toronto, Canada.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, The Netherlands: Foris.

Newport, E. L., Weiss, D. J., Wonnacott, E. & Aslin, R. N. (2004). Statistical learning in speech: Syllables or segments? *Paper presented at the 29th annual Boston University conference on language development*, Boston, MA.

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America, 102*(6), 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., & Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. In J. Cole & J. Hualde (Eds.), *Papers in laboratory phonology IX* (pp. 315–338). Berlin: Mouton de Gruyter.

Peperkamp, S., Le Calvez, R., Nadal, J.-P., & Dupoux, E. (2006a). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition, 101*(3), B31–B41.

Peperkamp, S., Skoruppa K., & Dupoux, E. (2006b). The role of phonetic naturalness in phonological rule acquisition. In: D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th annual Boston University conference on language development* (pp. 464–475). Sommerville, MA: Cascadilla Press.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 421–435.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology, 39*(3), 484–494.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-month-old Infants. *Science, 274*, 1926–1928.

Scott, D. R., & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior, 23*(4), 450–466.

Seidl, A., & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development, 1*(3–4), 289–316.

Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.

Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language, 25*, 169–201.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language, 51*(2), 173–189.

Singh, L., Morgan, J. L. & White, K. S. (in press). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development, 4* (2).

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York: Holt, Rinehart & Winston.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol. 2. Theoretical issues* (pp. 1157–1256). Hillsdale, NJ: Lawrence Erlbaum.

Soderstrom, M., White, K. S., Conwell, E., & Morgan, J. L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy, 12*, 1–29.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of statistical and stress cues to word boundaries by 7- and 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Trubetzkoy, N. S. (1939/1969). *Principles of phonology* (C. Baltaxe, Trans.). Berkeley: University of California Press (Original work published 1939).

Werker, J., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49–63.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition, 103*, 147–162.

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*(4), 501–528.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science, 30*, 945–982.