Phonological Representations in Language Acquisition:

Climbing the Ladder of Abstraction

Benjamin Munson

Department of Speech-Language-Hearing Sciences

University of Minnesota, Minneapolis

Jan Edwards

Department of Communicative Disorders and Sciences

University of Wisconsin, Madison

Mary E. Beckman

Department of Linguistics

Ohio State University, Columbus

1. Introduction

The physical world in which humans reside is limited to four dimensions, but the mental world in which our knowledge of language resides is not similarly limited. Individuals' knowledge of speech sounds comprises representations of information in multiple sensory domains, including representations of the auditory characteristics of the sounds that they have produced and have heard others producing, of the visual characteristics of the sounds they have seen others producing, and of the tactile, kinaesthetic, and somatosensory characteristics of sounds that they have produced.

This information is represented at multiple levels of abstraction in multiple domains of interpretation. Individuals who know English can interpret the duration of the interval of aperiodic energy between the release of a dorso-velar closure and the onset of voicing in a following front vowel as indexing a particular pattern of coordination among the gestures for the dorsal closure, the glottal opening, and the following vowel posture. This coordination pattern, in turn, can be interpreted as indicating the talker's intention to invoke a particular term in the system of paradigmatic phonation-type contrasts (a long voice-onset time indicates that the sound is part of the voiceless series of stop consonants) as well as a variety of syntagmatic facts about the utterance in which the syllable occurs, such as its meter (the /k/ is aspirated in *tomcat* but not in *bucket*) and its prosodic phrasing (the voice-onset time in /k/ is shorter in *tomcat* than in *Tom's cat*). The word-form [k^{h}æt] in an utterance of *Tom's cat* is indexed to the class of entities *felis catus*, just as the word-form [k^{h}ɪti] is in an utterance of the variant form *Tom's kitty*, and specific pronunciations of that word-form are indexed to attributes about the talkers who produced them. Details of the pronunciation of [æ], for example, can be interpreted as indexing a talker's age and sexual orientation (Johnson, 2006; Smith, Munson, and Hall, 2008), in perhaps a similar way to the way that the choice of *cat* over *kitty* can be interpreted as indexing the hearer's age and social relationship to the talker. Details of the pronunciation of the word form as a whole, similarly, may be interpretable as indexing whether the word form indexes the class of entities *felis catus* or a different class of entities: female *homo sapiens* including Tom's wife *Kat*, short for *Katherine*, as differentiated from Dick's sister *Kat*, short for *Katrina* (e.g., Jurafsky, Bell, & Girard, 2002; Gahl, 2008).

Thus, the categories that are indexed by an utterance of the word *cat* are of at least five types, including (1) the categories of intentions to posture the lips, tongue, glottis, etc., and to coordinate gestures for making different postures at the articulatory-more level, (2) the spectral patterns and auditory events that a listener parses to perceive the talker's articulatory intentions, (3) the terms in paradigmatic and syntagmatic contrast in the grammar of word-forms and phrases of the talker's and listener's shared language, (4) the classes of entities, properties and events that are indexed by particular word-forms and phrases, and (5) the types of social identities and relationships that are the larger cultural context for utterances that are produced by a talker and perceived by a listener who share a language,

This chapter discusses representations of the sound structure of language across the first part of the lifespan, from the time *in utero* at which the auditory system begins functioning, to early adulthood. Specifically, we focus on how phonological development involves building progressively more-abstract structures, starting with raw sensory encodings of the acoustic input that are first encountered in the womb, to the articulatory representations that begin in the first year of life, to the abstract representations that continue to develop throughout the lifespan. A more detailed theoretical justification for our approach can be found in Beckman, Munson, and Edwards (2007). Our proposal is similar to an independently developed one presented in

Pierrehumbert (2003). Our theoretical stance is that representations are latent variables. That is, representations are objects that can never be observed, but can only be inferred from individuals' overt behaviours. For example, consider aspects of a representation for the word *kitty*, which describes it as it comprising four phonemes grouped into two CV syllables, grouped into a trochaic foot. Positing such a representation could relate specific observable aspects of the pronunciation of this word, such as the aspiration of the word-initial /k/, or the realization of the medial /t/ as a flap, and the fact that hearing an utterance of this word can prime the production of phonologically related words like *city* or *kidder* or prosodically related words like *sofa*, to the observation that the word can be segmented into two syllables and four phonemes by many literate speakers of English. While the representation links together a wide variety of observable physical properties and behaviours, however, the representation itself is never observed or even observable. A useful analogy is that of the missing fundamental frequency. The perceived pitch of complex periodic sounds, like the glottal waveform, is influenced primarily by the frequency of the lowest harmonic, the fundamental frequency. However, when the fundamental frequency is removed, listeners still perceive the 'missing' fundamental as the spacing between the frequencies that remain. Representations are like the missing fundamental: we cannot observe them directly, but we can infer their existence from the behaviours that we do observe. This view is motivated in large part by observations about children's development, namely, that development involves children building phonological representations progressively as the consequence of producing and comprehending speech.

From this standpoint there are two mistakes to try to avoid in the discussion of representations of children's knowledge of speech sounds. One is to rely solely on observational methods, such as phonetic transcription, that inherently invoke models of fully-formed adult phonologies. The other is to over-interpret data from other types of observational methods in terms of representations that are motivated by accounts of adult behaviours.

2. Developmental changes in speech-sound knowledge

Building a model of phonological acquisition begins with the detailed study of children's knowledge of different aspects of sound structure of language, as well as of developmental changes in this knowledge. This section will review some of the basic facts about these developmental changes, calling specific attention to (a) the laboratory phonology methods that these studies have used, and (b) the implications of the findings for phonological representations, particularly as they relate to the five sources of categories described in our introduction. In reviewing these facts, we will be using the term *language-specific* to identify evidence that knowledge is phonological, and when we use this term, we will mean specific to a given ambient language, and not just specific to the capacity for human language in general. That is, for example, evidence that infants growing up in different speech communities behave differently is incontrovertible evidence of learning something about the language's sound structure, and is not just a development of the general capacity of primates (or of all animals) to produce and perceive vocal gestures for social purposes. There is evidence of language-specific categories at the lowest rungs of the "ladder of abstraction" well before there is evidence of knowledge of categories at the higher levels that are more typically associated with the term "phonological" in literature on the phonetics-phonology interface from the last century (see, e.g., Pierrehumbert, 1990; Keating, 1996). In light of the early evidence of language-specificity, an axiomatic approach to this interface is particularly counter-productive for understanding the development of phonological knowledge. That is, with Pierrehumbert, Beckman, and Ladd (2000, reprinted this volume), we think that the sources of "categoriality" are a proper object of study in their own right (see also Beckman & Edwards, 2000a). However, that debate is beyond the scope of this chapter, and we will use both "phonetic" and "phonological", as appropriate for the context, in referring to language-specific knowledge at lower levels of the ladder.

2.1. Perception

The development of knowledge of the sound structure of the ambient language begins very early. Infants begin to perceive speech in a language-specific way even before they are born, as evidenced by the finding that newborns prefer listening to the native language of their mothers as compared to another language with different prosodic characteristics (Nazzi, Bertoncini, & Mehler, 1998).

In the first few months of life, infants are also able to perceive many consonant contrasts (cf. Eimas, Miller, & Jusczyk, 1987, for a summary of this work). However, the contrasts they can perceive are not language-specific and, in fact, non-human animals can also perceive these same contrasts (e.g., Kuhl & Miller, 1978; Kuhl & Padden, 1983; Lotto, Holt, Kluender, 1997). The first strong evidence of language-specificity in perception of speech segments (as opposed to prosody) is observed at around six months when infants demonstrate a preference to listen to vowels that are more similar to native-language vowels (Kuhl et al., 1992) than ones that are not. By about nine months of age, infants stop attending to consonant contrasts that are not in their native language (e.g., Werker & Tees, 1984). This ability to home in on language-specific contrasts appears to be one of the first linguistic measures of how well an infant is learning the ambient language. Kuhl and colleagues (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005) found that English-speaking infants who were better able to discriminate a native consonant contrast (between /ta/ and /pa/) at 7.5 months had larger vocabularies at two years, relative to infants who were better able to discriminate a non-native contrast (between a Mandarin affricate and fricative). Houston and Jusczyk (2003) showed that infants as young as 7.5 month encode talker-specific information when listening to speech, such that infants were more accurate in recognizing words when produced by the same talker they had been familiarized with on a prior day.

As discussed in section 2.3, children's early productions may be subject to phonetic preferences. Two recent studies showed that individual differences in production preferences are mirrored in preferences in perception. Vihman and Nakai (2003) report that the strength with which infants prefer to listen to one contrast over the other is correlated with differences in the production of the same contrasts. Children preferred to listen to contrasts that they did not produce. DePaolis (2006) replicated this finding, but demonstrated that it was only operational in children with relatively more advanced vocal development. Children with less advanced development preferred to listen to contrasts that they preferred to produce. Together, Vihman and Nakai (2003) and DePaolis (2006) suggest a reciprocal relationship between early production preferences and attention in perception.

Toward the end of the first year of life, children begin to develop a receptive vocabulary. An influential study by Werker and Stager (2000) found that less phonetic detail is available to young children when a string of sounds can be interpreted as a label of an object than when simply listening to the same string. Fourteen month-old infants readily discriminate syllables such as [bɪ] versus [dɪ] when they are paired with complex visual displays such as checkerboard patterns, but do not discriminate them when they are paired with pictures of novel objects that the syllables could be labeling. This asymmetry is not shown by younger infants, suggesting that it does not emerge until children have learned a critical mass of vocabulary items. Subsequent work by Werker, Fennell, Corcoran, and Stager (2002) supports the suggestion that this asymmetry in discrimination is linked to vocabulary development, in that only those infants whose parents reported them producing at least 25 words (or comprehending at least 200 words) demonstrated the tendency. Those with smaller expressive or receptive vocabularies did not.

Beyond the first year of life, children’s speech perception abilities continue to develop. Much of this work been informed by the hypothesis that children’s phonological representations gradually become more fine-grained as they learn more and more words and must differentiate among them (e.g., Metsala & Walley, 1998). Hazan and Barrett (2000) tested 6- to 12-year-old children on acoustic continua associated with a variety of consonant contrasts and found that it was not until age 12 that children’s identification functions were as steep as adults. Clayards, Tanenhaus, Aslin, and Jacobs (2008) argued that the slope of identification funds reflects the shape of the experienced distributions of sounds being categorized. In this model, children's increasingly steep identification functions reflect their accrual of more sharply peaked distributions of sounds along different sensory parameters.

A number of studies have investigated differences in the weighting of acoustic cues by children and adults. In a series of studies, Nittrouer and colleagues (e.g., Nittrouer, 1992, 1996; Nittrouer & Miller, 1997; Nittrouer & Studdert-Kennedy, 1987) found that children's identification functions differed more strongly for /s/-vowel and /ʃ/-vowel sequences stimuli that differed in formant transitions than did adults'. Hence, children's fricative identification appear to rely more to formant transitions, while adults' judgments relate more to the fricative noise. Mayo, Scobbie, Hewlett, and Waters (2003) found that developmental differences in the influence of formant-transition and fricative-spectra cues for this same contrast were related to developmental changes in phonological awareness. Their seven-month longitudinal study showed that improvements in phoneme segmentation and blending (typically the most demanding of the tasks used to measure phonological awareness) occurred before shifts in cue weighting. Furthermore, phonological awareness, as measured at the earliest time periods of the study, predicted cue weighting strategies measured at the latest time period of the study.

Burnham (2003) reports a related finding for a measure of language-specific speech perception: the difference in degree of match to the sharp S-shaped curves of “categorical” identification functions between a native-language and very similar other-language continuum. This measure has a U-shaped curve in development, which peaks at around 6 years of age for English-speaking children, and is highly correlated with the same measures of phonological awareness that Mayo and colleagues used, as well as with measures of reading comprehension. This result provides further evidence of the protracted development of language-specific perception skills in the development of phonological categories at higher levels of abstraction that can be tapped in the acquisition of literacy in languages with alphabetic writing systems.

In sum, studies in this section underscore how perception both shapes and is shaped by the acquisition of language-specific phonological categories. Infants' prodigious early perception abilities allow them to encode the acoustic-phonetic detail needed to uncover the consonant and vowel categories relevant for the language being acquired. These in turn shape infants' subsequent perception of the same sounds in a manner that facilitates other aspects of linguistic processing, such as spoken-word recognition and word learning. At the same time, the extremely early onset of language-specificity in perception is an important reminder of the importance of looking at the input to the child.

2.2. Input

The field of language acquisition has come a long way from the days when the only thing that was said about the input was that children needed a little of it. Input to very young children (variously called “baby talk”, “motherese” and “infant-directed speech” or IDS) has now been studied in detail for more than thirty years. There are many characteristics of IDS that have been observed consistently across a range of studies in a number of languages using different methodologies. IDS relative to adult-directed speech (ADS) has a higher pitch, greater pitch range, shorter utterances, slower rate, and simpler syntax (e.g., Fernald & Simon, 1984; Fernald, Pinto, Swingley, Weinberg, & McRoberts 1998; Garnica, 1977; Lieven, 1994). One aspect of IDS that is particularly important for the establishment of phonological representation is the observation that speech sounds in IDS are hyper-articulated (e.g., Fernald, 2000) relative to the forms in ADS. However, this aspect of IDS is not found consistently across studies. While many studies (e.g., Kuhl et al., 1997, Liu, Kuhl, & Tsao, 2003) have found that IDS relative to ADS has more extreme formant frequencies for the point vowels /i, a, u/ in many languages, at least one language, Norwegian, does not show this pattern (Englund & Behne, 2006). Moreover, the evidence for hyper-articulation of specific consonant contrasts is inconsistent. For example, some studies have found a greater contrast in VOT between voiced and voiceless stop consonants in IDS relative to ADS (Malsheen, 1980), some studies have found a smaller contrast in IDS (Sundberg & Lacerda, 1999), and still other studies have found no difference (Baran et al., 1977). It has been suggested that hyper-articulation of consonant contrasts might be found in IDS to older infants, but not to younger infants, since older infants would be better able to make use of this information. In support of this claim, Sundberg and Lacerda (1999) found a smaller contrast in VOT for voiced and voiceless stops for IDS to 3-month-olds and a larger contrast in IDS to 11- to 14-month-olds. Cristiá (2009) found a similar result in her study of IDS to 4- to 6-month-olds and 12- to 14-month-olds. She examined the contrast between /s/ and /ʃ/ as measured by differences in the first spectral peak of the fricative spectrum and found a smaller contrast between the two fricatives in IDS to the younger group and a greater contrast in IDS to the older group.

Cristiá (2009) suggests that IDS may facilitate speech and language development either through its "affective" or its "informational" component. The affective component of IDS is signaled by its higher pitch and its greater pitch range, both of which are characteristic of positive emotion in speech (Scherer, 2003). Infants attend to IDS, at least in part, because they prefer to listen to speech with positive emotion. For example, they prefer "happy" ADS to neutral IDS (Singh, Morgan, & Best, 2002). The informational component of IDS relates to factors such as the shorter utterance durations, simpler syntax, and hyper-articulated vowel and consonantal contrasts, which may help infants learn linguistic contrasts. In support of this claim, Thiessen, Hill, and Saffran (2005) found that 6- to 8-month-old infants were able to segment speech into words that they had heard in IDS, but not in ADS. There are a small number of studies that demonstrate that individual differences in IDS predict individual differences in infant speech perception. Liu et al. (2003) found a positive correlation between the size of the vowel space in IDS to 6- to 8 and 10- to 12-month-old Mandarin-acquiring infants and these infants' ability to perceive the contrast between a Mandarin sibilant affricate and fricative. Cristiá (2009) looked at the same contrast in both IDS and infant perception. She found that 12- to 14-month-old infants who were better able to discriminate between /s/ and /ʃ/ had mothers who produced more distinct /s/ and /ʃ/ categories in their IDS.

In sum, the input that children receive during language acquisition provides them with at least some support for early vowel and consonant category formation, in that the signal may exaggerate the parameters that allow the child to uncover the categories in the signal itself. Much research continues to be needed in this area. Though Cristiá (2009) demonstrates a relationship between individual differences in category distinctiveness in adults' production and in the perception of children receiving this input, we know of no other study that reports this, nor any study reporting similar effects on children's subsequent productions.

2.3. Production

Children's vocal production changes dramatically in early development, perhaps most so during the first two years of life. In the first six months of life, children's productions progress from reflexive vocal behaviors (like crying and fussing) to sustained vocalizations suggesting independent control of respiration and phonation (Oller, 1980; Stark, 1980). Transitions in this early stage are likely driven by the growing autonomy of the different anatomic and physiologic systems used in speech production, as well as neural control of speech production that is separate from control of the same structures in non-speech tasks. Toward the middle of the first year of life, infants' vocalizations become both more varied and more speech-like. This phase culminates in children beginning the rhythmic articulatory movements of early canonical babbling. Phonetic transcription reveals strong consonant-vowel co-occurrence restrictions in this early babbling. These restrictions support an understanding of early babbling as a simple rhythmic wagging of the jaw, with different gross static tongue postures superimposed on the jaw cycle as a whole (MacNeilage & Davis, 1990). These co-occurrence constraints gradually relax. The gradual decoupling of consonant place and adjacent vowel quality reflects the infants' growing ability to control tongue movement separately from jaw movement within the jaw cycle.

Toward the end of the first year of life, infants begin to tailor their vocalizations to the characteristics of the language being acquired, as shown in the distribution of consonant places and manners of articulation in the babbled speech (de Boysson-Bardies, and Vihman, 1993), the formant patterns in vocoid portions (de Boysson-Bardies, Hallé, Sagart, & Durand, 1989), and the tempo and melody of the babble (Levitt, Utman, Aydelott, 1992; Whalen, Levitt, Wong, 1991). For example, Whalen et al. found that the babble of French-acquiring infants was more likely to contain the rising glissandi characteristic of adult French than was the babble of English-acquiring ones.

The phase in which children babble overlaps with that in which they produce their first words. As shown by Vihman, Macken, Miller, Simmons, and Miller (1985), the phonological characteristics of babble and first words is qualitatively very similar: children's production preferences in babbling correspond closely to their preferences in producing real words. Schwartz and Leonard (1982) found that children's early word learning is sensitive to production capacities, with children learning words that contain sounds over which they have productive control more readily than ones over which they do not. More recently, Storkel (2006) reported a similar relationship for children with larger-sized lexicons, demonstrating that children with less-rich phonological knowledge continue to restrict their word learning to forms that contain sounds over which they have productive control beyond the 50 word stage.

During the early stages of multisyllabic word and multi-syllable utterance production, languages' prosodic structure also influences production accuracy substantially. It is well documented that English metrical feet have a predominantly strong-weak structure (Cutler & Carter, 1987). Children's early productions often delete weak syllables so that the resulting productions fit this foot structure. This is true both in multisyllabic words (Gerken, 1994b) and multiword utterances (Gerken, 1994b).

Children's early words are coarse approximations of the adult forms. Transcriptions of toddlers' word productions are characterized by systematic errors such as deletions and substitutions relative to the target pronunciation. Transcription analyses of children's productions over the pre-school years typically show a rapid change in production patterns until they reach adult-like levels of accuracy for even the most challenging sounds. In large-scale normative studies of speech-sound acquisition in American English, this occurs by approximately age 6 (e.g., Smit et al., 1990).

Acoustic analyses of children's productions give a somewhat different picture of development. Detailed acoustic studies reveal that children's productions of sounds that are transcribed as substitutions for a target sound are often acoustically intermediate between and distinct from both the target sound and correct productions of the substitute. For example, Baum and McNutt (1990) showed that children's frontally misarticulated /s/ were distinct from their productions of /θ/, as well as other children's productions of /s/. Scobbie, Gibbon, Hardcastle, and Fletcher (2000) found that children's productions of target /st/ clusters, transcribed as /t/ or /d/, were acoustically distinct from correct productions of /t/ and /d/. Two recent studies found durational differences between productions that had an apparently deleted syllable or segment than ones that did not. Carter and Gerken (2004) Children's productions of verbs followed by a trisyllabic weak-strong-weak word with a deleted initial syllable were longer than those followed by a correctly produced strong-weak word. Song and Demuth showed that vowels were lengthened prior to deleted final consonants relative to productions without deletions. These findings and others suggest that phonological acquisition is continuous. Consistent with this, Li, Edwards, and Beckman (2009) demonstrated that Japanese- and English-acquiring children's production of anterior sibilant fricatives involves gradually greater acoustic differentiation across early phonological development. These

A parallel body of research has examined developmental changes in speech-sound duration and trial-to-trial variability in children's productions as a way of understanding developmental changes in the planning processes for speech. The logic underlying these studies is that longer duration and greater trial-to-trial variability indicates less-mature motor control. A well-established finding is that children's speech segment durations are longer than adults' (Eguchi & Hirsh, 1969; Kent & Forner, 1980, Smith, 1978, 1992; Smith & Kenney, 1994; Smith, Kenney & Hussein, 1996). This is true of children in all stages of phonological acquisition. Even after children have acquired perceptually accurate speech production, they continue to demonstrate longer speech-sound durations than adults. Kent and Forner (1980) showed that speech of children up to 12 years of age is measurably slower than that of adults. Many of these studies have also revealed that children have greater trial-to-trial variability in durational and spectral parameters than adults. Lee, Potamianos, and Narayanan (1999) showed that changes in variation in formant frequencies decreases with age throughout the teenage years. Munson (2004) showed that trial-to-trial variability in the spectral characteristics of /s/ decrease throughout the 3 to 8 year old age range.

Another set of studies has focused directly on articulatory movements and has examined developmental changes in kinematic parameters. Smith and Goffman (1998) examined trial-to-trial variability in lip movements in nonsense word productions. They found that children produce lip movements with greater trial-to-trial variability than adults. Goffman (1999) showed that stability in lip movements differentiated between children with different overall levels of linguistic development. Children with primary language impairments (i.e., language impairments that occur in the absence of a clear predisposing condition) produced nonwords with more variability than children with typical language development. Sasisekaran, Smith, Sadagopan, and Weber-Fox (2009) report that kinematic variability of children's nonword productions decreases over multiple days, suggesting that this variability reflects at least in part motor learning. This appears to be a hallmark of motor learning across the life-span, as shown by the finding that adults' nonword productions decrease in variability as well, provided the nonwords being repeated are sufficiently complex.

In sum, studies of children's speech production suggest that phonological development takes place over an extensive time period, not simply the first few years of life. It involves the acquisition of productions that are sufficiently adult-like as to be perceived and transcribed as accurate, and it also involves the development of adult-like speech-motor control. It is notable that the time course of development is appears to be considerably more protracted when production accuracy and motor control are assessed by acoustic analysis and direct kinematic measures, than when it is assessed by transcription alone. Hence, studies of the development of production abilities that focus solely on transcribed speech are likely to under-estimate the duration and complexity of this facet of phonological development. Developing tools to measure continuous change in children's speech is crucial to gain a fuller picture of phonological development. These tools include acoustic and kinematic analyses, as well as include novel auditory-perceptual rating scales, discussed in further detail in section 3.2.

2.4 Higher-level phonological knowledge

The studies reviewed in sections 2.2 and 2.3 document developmental changes in parametric phonetic knowledge of the distribution of sounds in two of the primary phonetic domains, articulation and acoustics. There is a parallel set of more-abstract representations that further categorize the speech signal. This section discusses the development of these representations.

The emergence of these more-abstract representations appears to be tightly yoked to the developmental expansion of the lexicon. Beckman and Edwards (2000b) and Edwards, Beckman, and Munson (2004) studied the relationship between the development of abstract phonological representations and vocabulary growth. These studies examined children's repetition of sequences of sounds embedded in nonwords. We reasoned that children could repeat a nonword-embedded sequence of sounds that occurs in many words, such as the sequence /ft/, by referencing the articulatory and acoustic representations for this sequence in known words in which it occurs, like *after* and *fifty*. In contrast, children's ability to repeat sequences of sounds that occur in few or no words, such as the sequence /fk/, cannot be made in reference to lexical knowledge. Repetition of these sequences would be supported by the existence of representations of objects like /f/ and /k/ that had been abstracted away from the sequences in which they occur. Consistent with this reasoning, Edwards, Beckman, and Munson (2004) found that the discrepancy between the accuracy of repetition of high- and low-frequency sequences (the *frequency effect*) decreased monotonically in children aged 3:0 (years:months) to 7:11. Further analyses found that the frequency effect was predicted statistically by measures of vocabulary size. Munson, Beckman, and Edwards (2005) showed that the predictive relationship between vocabulary size and the frequency effect held when developmental changes in real-word speech-production accuracy and speech perception were controlled statistically. That is, changes across development in the magnitude of the frequency effect appear to be distinct from developmental changes in parametric phonetic knowledge. Recent work by Zamuner (2009) partly replicated this finding and showed that the association is most robust for sounds in word-initial position.

These findings suggest that the emergence of abstract phonological representations in childhood is yoked to developmental changes in vocabulary size. One interpretation of the mechanism that underlies this association is that increases in vocabulary size lead to a reorganization of the lexicon along dimensions of phonological similarity. These dimensions become *de facto* representations of the sublexical units like phonemes and syllables. Beckman and Edwards (2000a) and Beckman, Munson, and Edwards (2007) hypothesized that there is a reciprocal relationship between the emergence of these representations and word learning. Their emergence allows children to interpret novel word-forms as combinations of known categories. This ability then allows children to form representations for novel strings more efficiently than if these strings were interpreted solely relative to existing articulatory and auditory representations. The fact that the emergence of abstract representations is yoked to developmental changes in vocabulary size gives a clue to *why* they emerge: one of their functions might be to allow the word learner to parse unfamiliar words as sequences of stored sublexical units during word learning. This in turn would facilitate the learning of new words. This conjecture would predict not only the relationship between lexical size and the integrity of sub-lexical units, but a relationship in turn between the integrity of these units and the trajectory of vocabulary development. A similar argument is made by Pierrehumbert (2003). This finding that has not yet been documented in longitudinal research. Cross-sectional studies of word learning have documented that children are more likely to learn high-phonotactic probability nonwords than low-probability ones (Storkel, 2001).

The hypothesis that developmental changes in abstract phonological knowledge are yoked to increases in vocabulary size is also consistent with Metsala and Walley's (1998) lexical restructuring hypothesis. The hypothesis posits that vocabulary growth drives developmental changes in children's overt judgments about the phonological structure of words, such as the ability to segment words into phonemes, and to blend phonemes produced in isolation into words. Indeed, the fact that such skills, known to be predictive of early reading ability, are predicted by vocabulary size suggests that the frequency effect in nonword repetition and the ability to make metaphonological judgments about the sound-structure of words both relate to the integrity of a coherent underlying set of sublexical representations.

Three themes can be extracted from this brief discussion. First, the development of higher-level phonological knowledge is a protracted process. Evidence can be found that these representations continue to change beyond the point at which children produce speech that is sufficiently adult-like to be transcribed as accurate. Second, multiple sources of evidence suggest that the development of these representations is yoked to developmental changes in vocabulary size. Third, higher-level phonological knowledge is multi-layered, and each of these layers is abstracted progressively further away from the parametric phonetic encodings.

3. Emerging areas of research

This section discusses new and promising approaches to the study of lexical and phonological representations in children. In the previous section, we reviewed evidence that phonological development involved detailed encodings in the primary sensory domains of speaking and listening, and higher-order generalizations about these encodings. The areas of research that we see as promising are those that expand on our knowledge of the encodings that children make and the generalizations that they impose on them. This includes cross-linguistic research and research on populations with atypical speech and language abilities. These expand on our knowledge base by examining a full range of variation in speech sounds and speech-sound knowledge than is possible in the study of typical speakers of only one language. We also see as promising studies that look at the different types of generalizations that children make over the parametric phonetic signals they have encoded, including those about the attributes of speakers who produce the speech they have heard. Finally, we regard as promising new techniques in measurement and analysis that allow us to better understand the speech children produce, and the cognitive architecture that allows them to make generalizations about it.

3.1 Rethinking Universals

The studies reviewed in sections 2.2, 2.3, and 2.4 demonstrate that development involves the accrual of knowledge in the primary sensory domains of audition, vision, proprioception, and somatosensory perception, as well as the development of progressively more abstract representations of this information. This section considers which aspects of these processes are universal, and which are language-specific. As discussed earlier, some language specificity is evident very early in development, even before the transition from babbling to first words. Productions become increasing language specific throughout early language development. Rvachew, Mattock, Polka, and Menard (2006) showed language-specific expansion of control of the vowel space in the second year of life in infants acquiring Canadian French or Canadian English. Buder and Stoel-Gammon (1994) found language-specific productions of /t/ in 2.5 year old children acquiring Swedish or English. Swedish children produced a more-diffuse spectrum for /t/ than English children, mirroring differences between the adult languages. Li, Edwards, and Beckman (2009) demonstrated language-specificity in two- and three-year-old children's productions of /s/, mirroring the cross-language differences seen in adults' productions of that sound.

Other recent studies have examined the claim, first articulated in Jakobson (1941), that the order in which children acquire speech sounds is relatively stable across languages and reflects universal constraints on the development and change of phonological systems. More recently, optimality theorists have made a similar claim, namely that all markedness constraints should outrank all faithfulness constraints in early child speech (e.g., Demuth, 1995). However, this claim was not supported by the results of Vihman and Velleman (2002) who examined spontaneous word productions of 20 children (five each for English, French, Japanese, and Welsh). They found that markedness constraints did not dominate faithfulness constraints, and, furthermore, they observed language-specific differences in the ranking of markedness and faithfulness constraints. In another study, in which the same word-initial consonant-vowel sequences were elicited across four languages, Edwards and Beckman (2008a) examined two- and three-year-old children's productions of lingual obstruents in Cantonese, English, Greek, and Japanese, and found substantial differences in the acquisition of what are ostensibly the same sounds across these languages. Multivariate statistical analyses showed that both language-specific constraints (specifically, the frequency of occurrence in the language) and language-universal constraints (presumably relating to universal ease of production and perception) were needed to account for the patterns across languages.

One example of a cross-linguistic asymmetry is in the acquisition of sibilant fricatives in English and Japanese. As reviewed by Beckman, Yoneyama, and Edwards (2003), large-scale studies of normal phonological development have found different orders of acquisition of /s/ and a corresponding post-alveolar sound (/ʃ/ in English, /ç/ in Japanese) in children acquiring English and ones acquiring Japanese. In English, /s/ is acquired before the postalveolar fricative, while in Japanese, it is acquired after it. Similar cross-linguistic asymmetries are found in the acquisition of /t/ and /k/: /t/ is acquired earlier than /k/ in English, and /k/ is acquired earlier than /t/ in Japanese. These findings run contrary to the assertion that coronals like /t/ and /s/ have a privileged status in phonological acquisition (e.g., Stemberger & Stoel-Gammon, 1991).

Clearly, there are substantial differences in phonological acquisition as a function of the language being acquired. What is less clear, however, is the extent to which the mechanisms that promote phonological category formation and abstraction across languages are similar. Some of the mechanisms described in section 2.2 seem likely candidates for universal applicability in at least spoken languages. Vocal-tract acoustics do not differ from language to language, and thus the nonlinear mapping between articulation and acoustics presumably enforces the same discretization of the phonetic space regardless of the (spoken) language being acquired. Moreover, the finding that infants tend to impute categories when given structured input is presumably a reflection of statistical learning abilities that are common to all humans. Maye, Werker, and Gerken (2002) and Maye, Weiss, and Aslin (2008) demonstrated that hearing infants can learn phonetic categories from non-random distributions of voice onset time (VOT) in stops. This learning would presumably occur regardless of the language being acquired. The finding also implies that if a contrast among phonological categories is to be learnable from the input, the shapes of the acoustic-phonetic distributions must reflect the category structure. Of course, the parameter(s) used by languages to convey categories are language specific. For example, VOT has different utility in characterizing voicing contrasts across languages, as demonstrated by Kong (2009), among many others. However, finding that VOT distributions do not mirror the category structure of the language does not necessarily mean that voicing cannot be learned from the distributions of acoustic parameters such as VOT; it might indicate instead that VOT is not the (sole) parameter whose distribution is useful for inducing voicing categories in that language.

Other mechanisms of abstraction might be candidates for other kinds of language-specificity. Consider Edwards, Beckman, and Munson's (2004) finding that phonotactic probability effects in nonword repetition are linked to developmental changes in vocabulary size. To our knowledge, this link has been explored in only one language other than English. Brea-

Spahn (2009) examined relationships between phonotactic probability, vocabulary size, and nonword repetition accuracy in a dialectally diverse group of children acquiring Latin American and Caribbean varieties of Spanish. Brea-Spahn created stimuli using a probabilistic phonotactic grammar similar to that used by Frisch, Large, and Pisoni (2000). In some ways, the behavior of participants in her experiment mirrored that seen previously: adults rated higher-probability stimuli as more like real Spanish words than lower-probability stimuli. Moreover, children repeated highly wordlike stimuli more accurately than less-wordlike stimuli. Given the correlation between wordlikeness and phonotactic probability, this comparison is analogous to the comparison between high- and low-probability stimuli by Edwards et al. However, groups that differed in vocabulary size, as assessed through a standardized test, did not differ in the magnitude of the wordlikeness effect on repetition. One possibility, suggested by Brea-Spahn, is that this negative result reflects the nature of the vocabulary test used in her study. That vocabulary test measures knowledge of words that are highly frequent in written Spanish. These words are more common across the diverse dialects being acquired by the children than are words that have high spoken frequencies. Perhaps a dialect-specific measure of individual children's spoken vocabulary size would predict the magnitude of the effect.

Another (admittedly more speculative) explanation of the lack of an interaction in Spanish is that the Spanish lexicon is less conducive to the kinds of generalization that are measured by phonotactic probability effects in the nonword repetition tasks of Edwards et al. and Brea-Spahn. For example, Spanish has words that are longer than those of English, and which are composed of simpler syllable shapes than those of English. These words are consequently less confusable with one another than English words, a finding that perhaps explains why phonological similarity has such different influences on lexical processing in Spanish and

English (Vitevitch & Rodriguez, 2005). A lexicon with longer, less confusable words might not exert the same pressures on the language learner to form representations with the same level of detail about the specific consonants and vowels that compose each different syllable as does the English lexicon, which comprises shorter words with more-complex syllable shapes that allow more types of minimal contrast, such as *cat* versus *scat* and *cats* as well as *cat* versus *pat* and *cap* or *cat* versus *coat* and *curt*. One can imagine even more striking differences between language pairs that differ even more than Spanish and English do.

The hypothesis that different lexicons lead to qualitatively different types of generalizations is partly supported by recent work by Beckman and Edwards (2008). Beckman and Edwards showed that the frequency of occurrence of different consonants in the ambient language lexicon predicts appreciably different proportions of variance in children's production accuracy of those consonants across a typologically diverse set of four languages. In two of the languages, English and Cantonese, consonant frequency predicts a substantial proportion of variance in children's word-initial consonant production accuracy. In the other two languages, Japanese and Greek, consonant frequency predicts much less variance in consonant production accuracy. English and Cantonese both have large vowel inventories and many monosyllabic words whereas Japanese and Greek both have only five vowels and very few monosyllabic words. These differences lead us to ask whether consonant segments extracted away from the following vowel context are the relevant type of representation for evaluating the relationships between category frequency and accurate production. An answer to this complex question can only come from large-scale cross-linguistic studies of speech-sound development that are informed by rigorous analyses of how languages carve the parametric phonetic space into categories, how these categories function in composing the words of the language, and how the earliest words are distributed along dimensions of similarity and contrast in the over-arching structural principles that organize the lexicon for rapid parsing and incorporation of new word forms. A number of promising ongoing research projects are examining this, including the παιδολογος project (Edwards & Beckman, 2008) and cross-linguistic work by Vihman and colleagues (e.g., Vihman, Thierry, Lum, Keren-Portnoy, & Martin, 2007).

3.2. Measurement within and across languages

A second emerging area in research regards the measurement of children's production accuracy and conformity to adult norms. As reviewed in section 2.2, transcription under-predicts patterns of variation in children's speech. Children's productions of what listeners perceive as identical productions are often acoustically distinct (Macken and Barton, 1980; Scobbie et al., 2000). However, there are many consonant contrasts for which no standard acoustic measure is yet available. Moreover, even when there is a well-developed acoustic measure (such as VOT for voicing contrasts), acoustic analysis also under-predicts cross-language differences in whether and when children's productions will be perceived as accurate by native speakers of the target language (as demonstrated, e.g., by Kong, 2009). Thus it is important to develop measures that can assess children's perceived speech production abilities in more robust detail than most transcription protocols allow.

Recent work with this aim has demonstrated that very subtle differences among types of productions can be assessed perceptually if the right measurement tools are used. For example, Schellinger, Edwards, Munson, and Beckman (2008) examined adults' perception of children's productions of /s/ and /θ/ which had been elicited via real-word and nonword-repetition tasks. Children's productions were carefully transcribed by a phonetically-trained native speaker of English, and tokens were chosen that the transcriber had coded as either [s], [θ], or something intermediate between [s] and [θ], to make for six different transcription categories: correct [s] for /s/ productions, [s] substitutions for target /θ/, productions coded as intermediate between [s] and [θ] but closer to [s], intermediate productions judged to be closer to [θ], [θ] substitutions for target /s/, and correct [θ] for /θ/ productions. A group of naïve native-English-speaking listeners were presented with these fricatives and were asked to rate how close they were to prototypical /s/ and /θ/ endpoints using a visual analog scaling (VAS) method. VAS was implemented by presenting listeners with a double-headed arrow bounded by the text "the 's' sound" and "the 'th' sound" and asking them to click at the location on the line that represented where they thought these sounds fell relative to the two endpoints. Quite strikingly, listeners' click locations discriminated among the same six different types of fricatives differentiated by the transcriber.

Other studies have shown that individual listeners' VAS judgments are well correlated with the acoustic parameters that differentiate between endpoints; that they are robust even to different levels of task difficulty (Kaiser, Munson, Li, Holliday, Beckman, Edwards, & Schellinger, 2009; Munson, Kaiser, & Urberg-Carlson, 2008); and that they are psychometrically superior to other continuous measures, such as direct magnitude estimates of phoneme goodness (Urberg-Carlson, Kaiser, & Munson, 2008). These measures hold great promise as ways to capture variation in children's productions that reflect a level of phonetic detail not captured by typical transcription protocols, and which better reflect the norms of a community of listeners, rather than the perceptions of a single transcriber.

Such work can also shed light on some of the apparent cross-linguistic differences discussed in section 3.1 (Edwards & Beckman, 2008a). For example, Munson, Li, Yoneyama, Hall, Beckman, Edwards, and Sunawatari (2008) examined whether the cross-language asymmetry in fricative acquisition between English and Japanese that Beckman et al. (2003)

found is attributable, in part, to differences in how adults in the ambient language environment interpret children's productions. Li et al found that Japanese-speaking adults accept a narrower range of children's productions (in a two-dimensional acoustic phonetic space) as acceptable tokens of /s/ than do English-speaking ones. This asymmetry suggests that language-specificity in how speakers parse the acoustic phonetic space is an additional source of cross-language differences in acquisition. It also emphasizes the importance of not relying solely on native-speaker transcribers' judgments when measuring phonological development.

3.3. Socioindexical learning

A third emerging area of inquiry is reflected in studies of children's acquisition of socially relevant variation in language. This work builds on concepts and methods in the growing subfield of sociophonetics, as summarized by Hay and Drager (2007), Foulkes (2008), and Docherty and Foulkes (2000), among others. This work has shown that individuals can mark their membership in different social groups through distinctive patterns of phonetic variation, and that listeners often use this variation to identify attributes about speakers. Much of this work has examined variation relative to macrosociological categories such as race, ethnicity, socio-economic status, and gender; however, more recent work has examined variation as it relates to local variation, such as social groups within schools (i.e., Drager, 2008). The influence of social variation on perception extends beyond tasks in which listeners make inferences about speaker attributes. Studies have shown that listeners' phonetic identification and spoken-word recognition can be biased by the social attributes they impute to the speakers they are listening to (Babel, 2009; Drager, 2006; Johnson, Strand, & D'Imperio, 1999; Munson, 2009; Niedzielski, 1999; Staum Casasanto, 2008; Strand, 2000; Strand & Johnson, 1996).

Understanding the acquisition of sociophonetic variation is important for at least two related reasons. First, learning the form-meaning mappings for social variants provides an additional level of complexity to the task of phonological acquisition. Given that lexical learning is one source of the richly-articulated hierarchy of multiple levels of category formation in regular phonological development, we would predict that sociophonetic learning might drive the development of an additional level of categorization of sounds. That is, when the child hears radically varying productions of ostensibly the same word form in the same prosodic context, and hears the same pattern of variation in another word form that shares a common subpart, the child is forced to abstract away some category structure for the common subpart that equates its variant forms. Second, just as abstract lexical phonological representations provide a scaffold for early word learners to parse systematic contextual variability in interpreting new strings of sounds, so might socioindexical stereotypes provide a scaffold for parsing phonetic variability that might otherwise appear to the learner to be random. That is, acquiring knowledge of the social categories allows the child to "parse out" the indexical function of the variation, to further promote unity of the different variants into one category at the cross-cutting level of abstraction from the lexicon. Put simply, socioindexical learning might impact the nature of phonological representations very directly and very significantly.

Children become aware of social variation in speech early in development, as illustrated by Patterson and Werker's (2002) finding that infants become sensitive to relationships between face gender and voice gender late in the first year of life. One large-scale study on the acquisition of socially relevant variation in production is provided by Docherty, Foulkes, Tillotson, and Watt (2006). Docherty et al. examined variation in medial /t/ variants in the variety of English spoken in Newcastle-upon-Tyne, England. Variants of /t/ in that dialect are stratified by sex, with women producing a pre-aspirated variant in medial position more often than men. Children demonstrate this sex asymmetry in production of this variant by about 3 ½ years. Smith, Durham, and Fortune (2007) examined the acquisition of standard and non-standard pronunciations of the MOUTH vowel (i.e., the vowel in the word *mouth*, using Wells' [1982] lexical-set notation) in a town in Northern Scotland. They found that children acquired the standard [ʌʉ] variant before the nonstandard [uː] variant, and that the rate of children's production of [uː] was correlated with the rate that the caregiver used it.

Recently, Li and Kong compared the acquisition of two gender-marked variants in a cross-sectional study of 2- to 5-year-old children (Li, Kong, Beckman, & Edwards 2008). The first is a variant of /ɕ/ in Mandarin spoken in Northern Mainland China. This variant is a feature of a style referred to as 'feminine accent'. It involves producing /ɕ/ with an especially high centroid frequency (Li, 2005; Li, Mays, Skorniakova, & Beckman, 2009). This affects the contrast among the three anterior sibilant fricatives (/s/, /ɕ/, /s̺/) of Mandarin in the two dimensions needed to differentiate them, centroid frequency and the frequency of the second formant of the following vowel at its onset. Li and Kong showed that the three-way contrast emerges in 2- and 3-year-old children, and is robustly present already in most of the older 3-year-old children. However, the gender-marked variant only occurs in the four- and five-year-old girls. That is, the socioindexical marker is acquired only after the regular phonological categories (i.e., those that are used to convey lexical contrast) are acquired. Li and Kong also examined children's production of a gender-marked voicing category in Japanese obstruents. As in Canadian French, Tokyo Japanese initial obstruents contrast a voiced series with a series that has voice-onset times intermediate between the short lag (voiceless) category of European French and the long lag (voiceless) category of English (e.g., Riney, Takagi, Ota, Uchida, 2007).

Moreover, one recent study (Kong 2009) reports that younger Tokyo Japanese women's VOT distributions for the former series are different from men's voiced stops in predominately showing tokens with no voicing lead. This shift of the VOT boundary between voiced and voiceless plosives is associated with a greater use of breathiness for voiceless stops in women's speech. Production of the more conservative "true" voiced stops conversely, has become a characteristic of men's speech. Li and Kong found that not even 5-year-old Japanese-acquiring boys had yet acquired this gender marker. However, true voiced stops in languages such as European French and Thai tend to be late acquired (see, e.g., Allen, 1985; Gandour, Petty Dardarananda, Dechongkit, & Mukongeoen, 1986). That is, the socioindexical marker had not yet been acquired, perhaps because it involves a regular phonological category that tends not to be adult-like until late in the preschool years.

One hypothesis that emerges from Li and Kong's work is that certain socioindexical markers might be acquired only after children have mastered production of the phonological categories themselves. This result in production is consistent with findings on the development of vowel perception in the first year of life. By about six months of age (Kuhl, Williams, & Meltzoff, 1991), infants prefer to look at images of faces whose postures match those required to produce the vowels that they are hearing (i.e., infants prefer to view an image of an adult with spread lips matched with the vowel /i/ than with the vowel /u/ or /ɑ/). However, at six months, have no preference for viewing an image of a male with spread lips and a male voice producing /i/, relative to that of an image of a female with spread lips and a male voice producing /i/. It is not until about ten months of age that infants show a preference for viewing an image of an adult talker matched with a voice of a talker of the same gender (Patterson & Werker, 2002). Crucially, the preference to match the face and voice gender emerges *after* the preference to match a facial posture and vowel quality emerges, at least at the group level. The time-course of regular phonological development and indexical development is clearly an important topic that is ripe for future studies.

3.4. Unified models of representations in children with disorders and typically developing children.

Models of typical language are made more powerful if they can also account for the abilities of individuals with atypical speech and language abilities, i.e., individuals with speech and language impairments. This section reviews selected studies on the phonological abilities of children with speech and language impairments in light of the models and studies discussed earlier. First, consider children with speech-sound disorder (SSD, sometimes referred to as phonological disorder [PD] or phonological impairment [PhI]). SSD is defined as significantly below age-level speech production in the absence of a clear medical or psychosocial etiology, such as hearing loss, intellectual impairment, structural anomaly, or a disorder of neuromotor control.

The error patterns that children with SSD make are often very systematic, and mirror those made by younger children acquiring the same language. Critically, children with SSD typically have appropriately-sized vocabularies for their age. Though some studies do report that children with SSD have slightly smaller-sized vocabularies than their peers without SSD, they are typically well above conventional cutoffs for language impairment. SSD provides an opportunity to understand the factors that contribute to variation in pronunciation while holding other factors known to affect pronunciation, like dialect and age, consistent. One consistent finding is that children with SSD have poorer speech perception ability than their age peers. Rvachew and Jamieson (1989) reported that children with SSD had less categorical perception of synthetic /s/-/ʃ/ and /s/-/θ/ continua than their typically developing peers. Edwards, Fox, and Rogers (2002) showed that children with SSD require more acoustic information than their peers to discriminate minimal pairs than typically developing age peers, and that their speech perception is proportionate to the severity of phonological impairment. Children with SSD also have poorer production motor control than their peers, as shown by Edwards's (1992) finding that they compensate more poorly for biomechanical perturbation than their peers with typical development. Hence, the perception difficulties of children with SSD appear to reduce their access to one important source of categories, distributions in the parametric phonetics. They also have reduced knowledge of articulatory-acoustic correspondences.

Three recent studies examined the relationship among measures of the speech-perception ability, speech production ability, and other measures of the phonological abilities of children with SSD. Munson, Edwards, and Beckman (2005) found that children with SSD do not differ from age-matched peers without SSD in the magnitude of the phonotactic probability effect in the nonword repetition task of Edwards, Beckman, and Munson, but did differ in the perception measure of Edwards, Fox, and Rogers. Again, measures of vocabulary size were better predictors of the magnitude of the phonotactic probability effect in nonword repetition than any other measure, including chronological age. Munson, Baylis, Kruase, and Yim (2006) found that speech-perception deficits better discriminated between children with and without SSD than did performance on two naming tasks that had been shown previously to index the robustness of lexical and phonological relationships. Rvachew and Grawburg (2006) analyzed relationships among speech perception, vocabulary size, and phonological awareness in a large cohort of children with and without speech-sound disorders. Rvachew and Grawburg measured children's receptive vocabulary size, their ability to discriminate between accurately and inaccurately produced tokens of sounds in real words, their speech-production accuracy, and their ability to make explicit judgments about the sound-structure of words (i.e., phonological awareness). Consistent with Munson et al.'s studies, Rvachew and Grawburg found that children's speech-production accuracy was predicted by their speech perception ability. They further found that speech perception predicted both vocabulary size and phonological awareness, and that vocabulary size predicted phonological awareness even after those two variables' relationship with speech perception was accounted for statistically. Hence, though there is ample evidence that children with SSD do not readily learn acoustic-phonetic distributions, they appear not to have specific difficulty in learning abstract phonological representations from their (presumably poorly phonetically specified) lexicon.

Second, consider children with primary language impairment (LI, sometimes referred to as Specific Language Impairment [SLI]). LI is defined as severe difficulty acquiring morphosyntactic, semantic, and lexical aspects of language in the absence of a clear predisposing condition. Though the primary presenting characteristics of children with LI are in aspects of language other than phonology, numerous research studies have shown that these children have difficulties in aspects of sound structure. Indeed, Bishop and Hayiou-Thomas (2008) suggested that these seemingly subtle difficulties may be the most clearly heritable aspects of LI. Children with LI have problems perceiving speech (Ziegler, Pech-Georgel, Alario, & Lorenzi, 2005), though the magnitude of this deficit is likely exaggerated by the use of synthetic speech stimuli in these experiments and the memory demands inherent in the tasks conventionally used to measure it (Coady, Evans, Mainela-Arnold, & Kleunder, 2007; Coady, Kleunder, & Evans, 2005). Moreover, children with LI produce speech with greater kinematic variability than typically developing children (Goffman, 2004). Munson, Kurtz, and Windsor (2005) reported that children with LI have larger phonotactic probability effects in nonword repetition than children with typical development. These were similar in magnitude to the effects in a younger group of children matched for vocabulary size, suggesting that the phonological representations of children with LI were no worse than would be predicted from the number of words they know. That is, children with LI have deficits not only in mapping the parametric phonetic space, but also in acquiring words. These lexical-acquisition difficulties lead to fewer opportunities to form abstract representations of the sound structure of language.

Together, studies of SSD and LI mirror many of the findings from many of the studies of typically developing children, in that they show that the primacy of speech perception in speech production, and the relationship between the size of the lexicon and the type of phonological generalizations that children make.

3.5 Computational Modeling

Models of normal functioning are made more powerful if they can be implemented computationally. Many existing models focus on learning of relationships between different aspects of the framework we have presented in this paper. For example, Plaut and Kello (1999) present a neural network model of the emergence of phonetic representations for sounds, modeling these as patterns of stable activation in a layer of hidden nodes mediating between acoustic representation and articulatory ones, that are sharpened by the link to semantic contrasts in the lexicon. Oudeyer (2005a) and Westermann and Miranda (2004) present two very different models of how articulatory-acoustic relationships might be learned. One model of the development of higher-level phonological knowledge (here, the learning of gradient and potentially violable constraints on the sound structure) is presented by Boersma and Hayes (2001).

In contrast, relatively few models have attempted to model the development of different types of phonological knowledge simultaneously. There are two notable recent exceptions. The first of these is Redford and Miikkulainen (2007), a model of the emergence of different syllable structures in development based on a combination of articulatory knowledge, perceptual knowledge, and the rate of lexical access. Oudeyer (2005b) models the development of one aspect of higher-level knowledge, phonotactic constraints, from a model of articulatory-acoustic associations. Again, this topic is ripe for future research, which will benefit from our increasingly sophisticated understand of parametric phonetics, acoustics, the lexicon, and the mechanisms of learning, as well as advances in the computational methods needed to develop viable models.

4. Conclusion

This chapter presented a framework for characterizing phonological knowledge along multiple sensory domains and at multiple levels of abstraction. Central to this framework is understanding the information in the primary sensory domains are encoded, and the factors that promote the emergence of higher-order representations that parse this variation into categories. We believe that this framework is useful in understanding how phonological representations change in development. Indeed, the studies that we reviewed suggest that phonological development can be understood as the gradual development of progressively more abstract structures in individuals' representation of language. Perhaps most importantly, though, we hope that the framework and findings reviewed in this chapter convince the reader that the status of something like /s/ as a category is enforced by multiple factors. Facts about the sound's articulation and acoustics, as well as how it functions in the lexicon and in socially situated communication all contribute to its cognitive representation. Moreover, there is ample evidence that adults' knowledge of these categories reflects their lifetime of producing and hearing sounds in words and in social communication. Researchers across disciplines should be cognizant of this when invoking categories like /s/ in other types of inquiry. The behavior of these categories—how they evolve during language change, how they are perceived, and how they are accessed—cannot be studied outside of considering how and why they arise, both in phylogeny and ontogeny.

References

Allen, G. (1985). How the young French child avoids the pre-voicing problem for word-initial voiced stops. *Journal of Child Language*, *12*, 37-46.

Babel M.E. (2009*). Phonetic and Social Selectivity in Speech Accommodation.* Unpublished Ph.D. Dissertation, Department of Linguistics, Berkeley, CA: University of California.

Baran, J.A., Zlatin Laufer, M., & Daniloff, R. (1977). Phonological contrastivity in conversation: A comparative study of Voice Onset Time*. Journal of Phonetics, 5*, 339-350.

Baum, S., & McNutt, J. (1990). An acoustic analysis of frontal misarticulation of /s/ in children. *Journal of Phonetics*, 18, 51-63.

Beckman, M. & Edwards, J. (2000a). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development, 71*, 240-249.

Beckman, M. & Edwards, J. (2000b). Lexical frequency effects on young children's imitative productions. In J. Pierrehumbert & M. Broe (eds.) *Papers in Laboratory Phonology* 5. Cambridge: Cambridge University Press.

Beckman, M., & Edwards, J. (2008). *Generalizing over Lexicons to Predict Consonant Mastery*. Oral presentation given at the Eleventh Conference on Laboratory Phonology, July 1, 2008, Wellington, New Zealand.

Beckman, M. E., Yoneyama, K., & Edwards, J. (2003) Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children. *Journal of the Phonetic Society of Japan*, 7, 18-28.

Beckman, M.E., Munson, B., & Edwards, J. (2007). The influence of vocabulary growth on developmental changes in types of phonological knowledge. In J. Cole & J. Hualde (Eds.), *Laboratory Phonology 9* (p. 241-264). New York: Mouton de Gruyter.

Bishop, D.V.M., & Hayiou-Thomas, M. (2008). Heritability of specific language impairment depends on diagnostic criteria. *Genes, Brain and Behavior, 7*, 365–372.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry, 32*, 45–86.

Brea-Spahn, M.R. (2009). *Spanish-Specific Patterns and Nonword Repetition Performance in English-Language Learners.* Unpublished Ph.D. Dissertation, University of South Florida.

Buder, E., & Stoel-Gammon, C. (1994). Cross-language differences in phonological acquisition: Swedish and American /t/. *Phonetica, 51*, 146-58.

Burnham, D. (2003). Language specific speech perception and the onset of reading. *Reading and Writing: An Interdisciplinary Journal*, 16**:** 573-609.

Clayards, M. Tanenhaus, M.K., Aslin, R., & Jacobs, R.A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition, 108*, 804-809.

Coady, J. A., Kluender, K. R., & Evans, J. L. (2005). Categorical perception of speech by children with specific language impairments. *Journal of Speech, Language, and Hearing Research*, 48, 944-959.

Coady, J. A., Evans, J. L., Mainela-Arnold, E., & Kluender, K. R. (2007). Children with specific language impairments perceive speech most categorically when tokens are natural and meaningful. *Journal of Speech, Language, and Hearing Research, 50*, 41-57.

Cristia, A. (2009). *Individual variation in infant speech processing: Implications for language acquisition theories.* Unpublished doctoral dissertation, Purdue University.

Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English Vocabulary. *Computer Speech and Language, 2*, 133-142.

de Boysson-Bardies, B., Hallé, P., Sagart, L., & Durand, C. (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language*, *16*, 1-17.

de Boysson-Bardies, B., & Vihman, M.M. (1991). Adaptation to language: babbling and first words in four languages. *Language, 67*, 297-319

Demuth, K. (1995). Markedness and the development of prosodic structure. In J. Beckman (Ed.) *Processings of the North Eastern Linguistic Society 25* (pp. 13-25). Amherst, MA: GLSA, University of MA.

DePaolis, R. A. (2006). The influence of production on the perception of speech. In Bamman, D., Magnitskaia, T. & Zaller, C. (eds.), *Proceedings of the 30th Boston University Conference on Language Development* (p. 142-153). Somerville, MA: Cascadilla Press.

Docherty G.J., & Foulkes P. (2000). Speaker, speech, and knowledge of sounds. In: N. Burton-Roberts, P. Carr & G.J. Docherty (Eds,) *Phonological knowledge: conceptual and empirical issues* (p. 105-129).. Oxford, Oxford University Press.

Docherty, G. J., Foulkes, P., Tillotson, J., & Watt, D. (2006). On the scope of phonological learning: issues arising from socially structured variation. In L. Goldstein, D.H. Whalen, and C. Best, (Eds.). *Laboratory Phonology 8* (p. 393-421). New York: Mouton de Gruyter.

Drager, K. (2006). From Bad to Bed: the relationship between perceived age and vowel perception in New Zealand English. *Te Reo, 48*, 55-68 .

Drager, K. (2008). *Sensitivity to Grammatical and Sociophonetic Variability in Perception.* Oral presentation given at the Eleventh Conference on Laboratory Phonology, July 1, 2008, Wellington, New Zealand.

Edwards, J., & Beckman, M.E. (2008a). Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language Learning and Development, 4*, 122–156.

Edwards, J., & Beckman, M. E. (2008b). Methodological questions in studying phonological acquisition. *Clinical Linguistics and Phonetics*, 22, 939-958.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*, 421-436.

Edwards, J. (1992). Compensatory speech motor abilities in normal and phonologically disordered children. *Journal of Phonetics 20*, 189-207.

Edwards, J., Fox, R. A., & Rogers, C. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and phonetic inventory size. *Journal of Speech, Language, and Hearing Research*, *45,* 231-242.

Eguchi, S., & Hirsch, I. (1969). Development of speech sounds in children. *Acta Otolaryngologica Supplementum, 257*, 1-51.

Eimas, P. D., Miller, J. L., & Jusczyk, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.) Categorical perception: The groundwork of cognition (pp. 161-195). New York: Cambridge University Press.

Englund, K., & Behne, D. (2006). Changes in infant directed speech in the first six months. *Infant and Child Development, 15*, 139-160.

Fernald, A. (2000). Speech to infants as hyperspeech: Knowledge-driven process in early word recognitions. Phonetica, 57, 242-254.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. Developmental Psychology, 20, 104-113.

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, 228-231.

Foulkes, P. (2008). *Exploring social-indexical knowledge: a long past but a short history.* Oral presentation given at the Eleventh Conference on Laboratory Phonology, July 1, 2008, Wellington, New Zealand.

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on processing of non-word sound patterns. *Journal of Memory and Language, 42*, 481-496.

Gahl, S. (2008.) "Thyme" and "Time" are not homophones. Word durations in spontaneous speech. *Language, 84*, 474-496.

Gandour, J., Petty, S. H., Dardarananda, R. , Dechongkit, S., & x Mukongoen, S. (1986). The acquisition of the voicing contrast in Thai: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *13*. 561-572.

Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C. Gallaway & B. J. Richards (Eds.), *Talking to children: Language input and acquisition*. New York: Cambridge University Press.

Gerken, L. (1994a). Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. *Journal of Memory and Language, 33*, 19-38.

Gerken, L. (1994b). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, *21*, 565-584.

Goffman, L. (1999). Prosodic influences on speech production in children with specific language impairments and speech deficits: kinematic, acoustic, and transcription evidence. *Journal of Speech, Language, and Hearing Research, 42*, 1499-1517.

Goffman, L. (2004). Kinematic differentiation of prosodic categories in normal and disordered language development. *Journal of Speech, Language, and Hearing Research, 47*, 1088-1102.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetics, 28*, 377-396.

Hay, J., & Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology, 36*, 89-103.

Hazan, V. & Barrett, S. (2000). The development of phonetic categorization in children aged 6-12. *Journal of Phonetics, 4*, 377-396.

Hoover, J.R. & Storkel, H.L. (2007). Phonological and lexical cues in word learning by preschool children. in seminar entitled *Word Learning in Situ: Interplay between Learners and Learning Environments* (Convener: K.K. McGregor). American Speech-Language-Hearing Association Convention, Boston, MA.

Houston, D., & Jusczyk, P. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1143-1154.

Jakobson, R. (1941). *Child Language, Aphasia and Phonological Universals.* The Hague: Mouton de Gruyter.

Johnson, S. (2006). *If You're 'heppy' and You Know it, Front Your /æ/.* Oral presentation given at NWAV 36, Columbus, OH.

Johnson, K, Strand, E. & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, 27, 359-384.

Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In C. Gussenhoven & N Warner (eds.), *Papers in Laboratory Phonology VII*. Berlin/New York: Mouton de Gruyter, 1-34.

Kaiser, E., Munson, B., Li, F., Holliday, J., Beckman, M.E., Edwards, J., & Schellinger, S. (2009). Why do adults vary in how categorically they rate the accuracy of children's speech? *Journal of the Acoustical Society of America, 125*, 2753. Downloaded on June 14, 2009 from http://www.ling.ohio-state.edu/~edwards/ASA09_Kaiser_etal_poster.pdf.

Keating, P. A. (1996). The phonetics-phonology interface. *Studia Grammatica*, 41, 262-278.

Kent, R., & Forner, L. (1980). Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics, 8*, 157-168.

Kong, E.J. (2009). *The Development of Phonation-type Contrasts in Plosives: Cross-linguistic Perspectives.* Unpublished Ph.D. Dissertation, Department of Linguistics. Columbus, OH: Ohio State University.

Kuhl, P. K. & Miller, J. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America, 63*, 905-917.

Kuhl, P. K. & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature for macaques. *Journal of the Acoustical Society of America, 71*, 1003-1010.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Kozhnevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684-686.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the critical period. *Language Learning and Development, 1*, 237-264.

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.

Lee, S., Potamianos, A. and Naryanan, S. (1999). Acoustics of children's speech: developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*, 1455-1468.

Levitt, A., & Utman, J.., (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language, 19*, 19-49.

Li, F.(2005) The Production and Perception of Dental vs. Retroflex Sibilants in the Songyuan Dialect of Northeastern Mandarin Chinese. Poster presented in NWAV 35, November, 2005, New York.

Li, F., Kong, E.J., Beckman, M.E., & Edwards, J.(2008). *Adult Acoustics and Developmental Patterns for Gender-marked Phonetic Variants in Mandarin Fricatives and Japanese Stops,* Poster presented at 11th Laboratory Phonology, Wellington, New Zealand, June 30, 2008.

Li, F., Edwards, J., and Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics, 37*, 111–124.

Li, F., Mays, C., Skorniakova, O., & Beckman, M.E. (2009). *Gendered production of sibilants in the Songyuan dialect of Mandarin Chinese.* Poster presented at the 2009 Annual Meeting of the Linguistic Society of America, San Francisco, 8-11 January 2009.

Lieven, E. V. M. (1994). Crosslinguistic and cross cultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (Eds.), *Input and Interaction in Language Acquisition* (pp. 56-73). Cambridge: Cambridge University Press.

Lotto, A. J., Kluender, K. R., & Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (Coturnix coturnix japonica), *Journal of the Acoustical Society of America*, *102 (2)*, 1134-1140.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M., (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6, F1-F10.

Macken, M. A. and D. Barton .(1980). The Acquisition of the Voicing Contrast in English: A Study of Voice Onset Time in Word-Initial Stop Consonants. *Journal of Child Language,* 7, 41-74.

MacNeilage , P. F., & Davis, B. L. (1990). Motor explanations of babbling and early speech patterns**.** In M. Jeannerod (Ed.), *Attention and performance XIII: motor representation and control*, pp. 567–582. Hillsdale, NJ: Lawrence Erlbaum.

Malsheen, B. J. (1980). Two hypotheses for phonetic clarification in the speech of mothers to children. In G. H Yeni-Komshian, J. F. Kavanaugh, & C. A. Ferguson (Eds.), *Child Phonology* (Vol. 2, pp. 173-184). San Diego, CA: Academic press.

Maye, J., Weiss, D., & Aslin, R. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science, 11*, 122-134.

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

Mayo, C., Scobbie, J., Hewlett, N., & Waters., D. (2003). The influence of phonemic awareness development on acoustic cue weighting in children's speech perception. *Journal of Speech, Language and Hearing Research, 46*, 1184-1196.

Metsala, J., & Walley, A. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J.L. Metsala & L.C. Ehri (Eds.), *Word Recognition in Beginning Literacy*, (p. 89-120). Mahwah, NJ: Erlbaum,

Munson, B. (2004). Variability in /s/ production in children and adults: evidence from dynamic measures of spectral mean. *Journal of Speech, Language, and Hearing Research, 47*, 58-69.

Munson, B. (2009). *Gender Biases in Fricative Perception, Revisited.* Oral presentation given at the 2009 meeting of the Linguistic Society of America, San Francisco, CA. Downloaded on June 13, 2009 from http://www.tc.umn.edu/~munso005/LSA2009_Munson.pdf.

Munson, B., Baylis, A.L., Krause, M.O., & Yim, D.-S. (in press). Representation and Access in Phonological Impairment. In C. Fougeron, M. D'Imperio, N. Vallee, & B. Kuhnert (Eds.), *Laboratory Phonology 10*. New York: Mouton de Gruyter. Downloaded on June 14, 2009 from http://www.tc.umn.edu/~munso005/MunsonEtAl_LabPhon10_InPress.pdf.

Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research, 48*, 61-78.

Munson, B., Edwards, J., Schellinger, S., Beckman, M.E., & Meyer, M.K. (2009). Deconstructing phonetic transcription: covert contrast, perceptual bias, and an extraterrestrial view of *Vox Humana*. Accepted pending revisions in *Clinical Linguistics and Phonetics*.

Munson, B., Kaiser, E., & Urberg Carlson, K. (2008). *Assessment of children's speech production 3: Fidelity of responses under different levels of task delay*. Poster presented at the 2008 ASHA Convention, Chicago, 20-22. Downloaded on June 14, 2009 from http://www.tc.umn.edu/~munso005/MunsonKaiserUrberg-Carlson_Final.pdf.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*, 1033-1047.

Munson, B., Li, F., Yoneyama, K., Hall, K., Beckman, M., Edwards, J., & Sunawatari, Y. (2008). *Sibilant Fricatives in English and Japanese: Different in Production or in Perception?* Oral presentation given at the annual meeting of the Linguistic Society of America, Chicago, IL

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 756-766.

Niedielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18, 62-85.

Nittrouer, S. & Miller, M.E. (1997). Predicting developmental shifts in perceptual weighting schemes, *Journal of the Acoustical Society of America, 101*, 2253-2266.

Nittrouer, S. & Studdert-Kennedy, M. (1987). The role of coarticulatory effects on the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30*, 319-329.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics, 20*, 351-382.

Nittrouer, S. (1996). The discriminability and perceptual weighting of some acoustic cues to speech perception by three-year-olds. *Journal of Speech and Hearing Research, 39*, 278-297.

Oller, R.K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child Phonology I: Production*. New York: Academic Press.

Oudeyer, P.-Y. (2005a) The Self-Organization of Speech Sounds, *Journal of Theoretical Biology*, *233,* 435-449.

Oudeyer, P.-Y. (2005b) How phonological structures can be culturally selected for learnability, *Adaptive Behavior, 13*, 269-280.

Patterson, M., & Werker, J.F. (2002). Infants' ability to match dynamic information in the face and voice. *Journal of Experimental Child Psychology, 81*, 93-115.

Pierrehumbert, J. (1990). Phonological and Phonetic Representation. *Journal of Phonetics*, 18, 375-394.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2-3), 115-154.

Pierrehumbert, J., Beckman, M. E., & Ladd, D. R. (2000). Conceptual foundations of phonology as a laboratory science. In Noel Burton-Roberts, Philip Carr, & Gerard Docherty, eds., *Phonological Knowledge: Its Nature and Status*, pp. 273-303. Oxford University Press. [Reprinted, this volume.]

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: a distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381-415). Mahwah, NJ: Elrbaum.

Redford, M.A., Miikkulainen, R. (2007). Rate effects on structure in a source-filter model of phonological development. Language, 83, 737-69.

Riney, T., Takagi, N., Ota, K., & Uchida. Y. (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics*, *35*. 439-443.

Rvachew, S., & Grawburg, M. (2006). Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 49*, 74-87

Rvachew, S., Mattock, K., Polka, L., & Menard, L. (2006). Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *Journal of the Acoustical Society of America, 120*, 2250-2259

Rvachew, S. & Jamieson (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders 54*, 193-208.

Sasisekaran, J., Smith, A., Sadagopan, J., & Weber-Fox, C. (in press). *Nonword Repetition in Children and Adults: Effects on Movement Coordination. Developmental Science*.

Schellinger, S. K., Edwards, J., Munson, B., & Beckman, M. E. (2008). *The role of listener expectations on judgments of children's /s/ productions.* Poster presented at the Symposium on Research in Child Language Disorders, 5-7 June, University of Wisconsin, Madison.. Downloaded June 17, 2009 from http://www.ling.ohio-state.edu/~edwards/SRCLD_schellinger_final.pdf.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227-256

Schwartz, R., & Leonard, L. (1982). Do children pick and choose? An examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language, 9*, 319-336.

Scobbie, J., Gibbon, F., Hardcastle, W., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (p. 194-207). Cambridge, MA: Cambridge University Press.

Singh, L., Morgan, J., & Best, C. T. (2002). Baby talk or happy talk? *Infancy, 3*, 365-394.

Smit, A. B., Freilinger, J. J., Bernthal, J. E., Hand, L., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779-798.

Smith, A., & Goffman, L. (1998). Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language, and Hearing Research, 41*, 18-30.

Smith, B. (1978). Temporal aspects of English speech production: a developmental perspective. *Journal of Phonetics, 6*, 37-67.

Smith, B. (1992). Relationships between duration and temporal variability in children's speech. *Journal of the Acoustical Society of America 91*, 2165-2174.

Smith, B. & Kenney, M.K. (1994). Variability control in speech production tasks performed by adults and children. *Journal of the Acoustical Society of America 96*, 699-705.

Smith, B. Kenney, M.K. & Hussain (1996). A longitudinal investigation of duration and temporal variability in children's speech production. *Journal of the Acoustical Society of America, 99*, 2344-2349.

Smith, J., Durham, M., & Fortune, L. (2007). "Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change, 19*, 63-99.

Staum Casasanto, L. (2008). *Experimental Investigations of Sociolinguistic Knowledge.* Unpublished Ph.D. Dissertation, Department of Linguistics. Palo Alto, CA: Stanford University.

Smith, E.A., Munson, B., & Hall, K.C. (2008). *Rethinking the meaning of Minnesotan [æ]: sexual orientation or personal well being?* Oral presentation at the conference on New Ways of Analyzing Variation (NWAV), Houston, TX. Downloaded June 17, 2009 from http://www.ling.ohio-state.edu/~kchall/Smith_Munson_Hall_NWAV_2008.pdf.

Song, J.Y., & Demuth, K. (2008_. Compensatory vowel lengthening for omitted coda consonants: A phonetic investigation of children's early representations of prosodic words. *Language & Speech, 51*, 382-399.

Stark, R. (1980).Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child Phonology I: Production*. New York: Academic Press.

Storkel, H. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44*, 1321-1338.

Storkel, H.L. (2006). Do children still pick and choose? The relationship between phonological knowledge and lexical acquisition beyond 50 words. *Clinical Linguistics & Phonetics, 20*,, 523-529.

Strand, E. (2000). *Gender Stereotype Effects in Speech Processing.* Ph.D. Dissertation, Department of Linguistics. Columbus, OH: The Ohio State University.

Strand, E., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference, Bielfelt, October 1996* (pp. 14-26). Berlin: Mouton de Gruyter.

Sundberg, U., & Lacerda, F. (1999). Voice Onset Time in speech to infants and adults. Phonetica, 56, 186-199.

Thiessen, E. D., Hill, E. & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. Infancy, 7, 53-71.

Urberg Carlson, K., Kaiser, E., & Munson, B. (2008). *Assessment of children's speech production 2: Testing gradient measures of children's productions.* Poster presented at the 2008 ASHA Convention, Chicago, 20-22. Downloaded on June 14, 2009 from http://www.tc.umn.edu/~munso005/Urberg-CarlsonEtAl_Final.pdf.

Vihman, M. M. & Velleman, S. (2002). The optimal initial state. Unpublished manuscript available on the Rutgers Optimality Archive, http://roa.rutgers.edu/index.php3.

Vihman, M. M. & Nakai, S. (2003). Experimental evidence for an effect of vocal experience on infant speech perception. In M. J. Solé, D. Recasens & J. Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona (pp. 1017-1020).

Vihman, M., Macken, M., Miller, R. Simmons, H., & Miller, J. (1985). From babbling to speech: a reassessment of the continuity issue. *Language 61*, 397-445.

Vitevitch, M.S. & Rodríguez, E. (2005). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders, 3,* 64-73.

Wells, J. (1982). *Accents of English, Volume 1*. Cambridge: Cambridge University Press.

Werker, J.F., & Stager, C.L. (2000). Developmental changes in infant speech perception and early word learning: Is there a link? In M. Broe & J.B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology 5* (pp.181-193). Cambridge University Press.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, ***3***, 1-30.

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

Westermann, G. and Miranda, E. (2004) A new model of sensorimotor coupling in the development of speech. *Brain and Language, 89*, 393-400

Whalen, D.H., Levitt, A., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants, *Journal of Child Language, 18,* 501–516

Zamuner, T.S. (2009). Phonological probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research, 52*, 49-60.

Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F.-X., & Forenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences, 102*, 14410-14115.