Understanding the developing sound system: interactions between sounds and words

Suzanne Curtin¹* and Tania S. Zamuner²

Over the course of the first 2 years of life, infants are learning a great deal about the sound system of their native language. Acquiring the sound system requires the infant to learn about sounds and their distributions, sound combinations, and prosodic information, such as syllables, rhythm, and stress. These aspects of the phonological system are being learned simultaneously as the infant experiences the language around him or her. What binds all of the phonological units is the context in which they occur, namely, words. In this review, we explore the development of phonetics and phonology by showcasing the interactive nature of the developing lexicon and sound system with a focus on perception. We first review seminal research in the foundations of phonological development. We then discuss early word recognition and learning followed by a discussion of phonological and lexical representations. We conclude by discussing the interactive nature of lexical and phonological representations and highlight some further directions for exploring the developing sound system. © 2014 John Wiley & Sons, Ltd.

INTRODUCTION

Learning the phonological system of the language requires the infant to learn about sounds and their distributions, sound combinations, and prosodic information, such as syllables, rhythm, and stress. These aspects of the phonological system are being learned simultaneously as the infant experiences the language around him or her. What binds all of the phonological units is the context in which they occur, namely, words. Recent work demonstrates that even very young infants are beginning to acquire some of the words in their native language. In illustration, Tincoff and Jusczyk¹ and Bergelson and Swingley² found that infants as young as 6 months can recognize common words, and Bortfeld et al.³ demonstrated that highly frequent forms, such as mommy and the child's name can facilitate segmentation of adjacent items by infants as young as 6 months. At the same time, infants begin segmenting word forms using various other available cues, suggesting that early word-form representations are being stored and might therefore be available for subsequent processing of the input (for a review see Ref 4). Thus, in our review exploring the development of phonetics and phonology, we showcase the interactive nature of the developing lexicon and sound system. To this end, we first review seminal research in the foundations of phonological development with a focus on perception. We then discuss early word recognition and learning, followed by a discussion of phonological and lexical representations. We conclude by discussing the interactive nature of lexical and phonological representations and highlight some further directions for exploring the developing sound system.
FOUNDATIONS IN PHONOLOGICAL DEVELOPMENT

Speech Perception

A number of seminal studies have explored what information in the speech signal infants prefer to listen to and what they can discriminate. It has been shown that very young infants prefer listening to speech over nonspeech, such as white noise and sine-wave analogs of speech. From birth, infants demonstrate a preference for their native language’s rhythm pattern. Infants under 6 months can discriminate trochaic (Sw) and iambic (wS) stress patterns. By 9 months, English-learning infants prefer strong–weak (trochaic) stress patterns to weak–strong (iambic). It is at this age that we see evidence of infants preferring sound sequences (i.e., phonotactics) that are legal and more frequent in their native language, and preferences for sequences with high type phonotactic probabilities (PPs). Thus, there is abundant evidence that infants are learning about the phonological system over the course of the first year of life, since they demonstrate preferences for phonological patterns that occur in the ambient environment, including sound combinations and prosodic patterns. Infants’ sensitivity to various cues in the language input can facilitate the segmentation of words from the continuous stream of speech and this bidirectional relationship can mutually support the building of the lexicon and emergence of the phonological system.

Seminal research in infant speech perception suggested that infants are born as universal listeners. Indeed, it was posited that infants discriminate speech sound contrasts found in all of the worlds’ languages from birth. Building on that, important research demonstrated that infants begin to hone in on their native-language vowel categories around 6–8 months of age, and that they begin to focus in on their native-language consonant categories around 10-12 months of age. Not only do infants appear to attend to those contrasts in their native language, but also they show improved discrimination of some native-language speech sound contrasts, such as /d/ and /b/. Moreover, nonnative contrasts that fall outside of an infant’s native-language sound categories, such as clicks which occur in languages such as Xhosa and Zulu, continue to be discriminated by infants and adults. This body of research, much like the research exploring native-language prosody and phonotactics, provides further support for early phonological development, at least in terms of focusing on speech sound distinctions.

Subsequent research has found that not all speech sound contrasts are equally discriminable. In some cases, it is because of asymmetrical processing and in others it is because of the salience of the contrast. Some of the first work to demonstrate this was by Kuhl who posited that extreme productions of vowels in perceptual space act as magnets by attracting less prototypical tokens into that categorical space. Polka and Bohn further found that an infant’s perception of a vowel may shift the perception of a subsequent vowel. This phenomenon relates to the prominence or importance of point (extreme) vowels (e.g., /i/, /u/, and /a/). That is, when infants are first presented with one of point vowels (e.g., /i/) and then hear, a more central vowel (e.g., /I/) the central vowel appears to be encompassed into the point vowel’s auditory space, making it difficult for infants to discriminate between them. In the case of consonants, it has been shown that acoustic salience influences early speech sound discrimination. Narayan et al. have shown that perceptually difficult contrasts require more experience in order for infants to discriminate between them. For example, Filipino and English infants both have difficulty in discriminating the alveolar /na/ versus the velar /nas/ nasal stop contrast in word-initial position, even though the contrast exists in this position in Filipino, and in syllable final positions in English. It is not until later in development, around 10 months, that Filipino-learning infants discriminate this contrast, whereas English-learning infants at the same age do not discriminate this contrast. Difficulty in discriminating contrasts has also been found for other native-language contrasts in noninitial positions. In word-final (coda) positions, Dutch-learning infants less than 16 months of age do not discriminate voiceless stops contrasts that differ in their place of articulation (POA). There is some emerging evidence that English-learning infants at 12 months can discriminate POA contrasts in word-final codas, but only when the speech sounds are voiced (e.g., ab vs ag), as they fail to discriminate voiceless speech sounds in this position (e.g., ap vs ak). These findings suggest that experience and salience of the contrast are integral to infants’ ability to discriminate speech sounds.

What other information might facilitate discrimination and the learning of sound contrasts? Some of the possibilities include statistical learning mechanisms, the visual input, and the specific context in which the contrast occurs. For example, infants of 8 months can use visual information to track monomodal versus bimodal distributions of speech sounds and this impacts subsequent perception of speech sounds.
a unimodal distribution of a continuum between /ba/ and /da/ while simultaneously presented with synchronous visual articulations of both a canonical /ba/ and /da/, they could discriminate the contrast. However, when a single visual articulation of either /ba/ or /da/ was paired with the monomodal continuum, infants failed to discriminate the contrast.30 Yeung and Werker31 found that while English-learning infants no longer show discrimination of nonnative contrasts, such as the Hindi dental/retroflex contrast,19 in a purely acoustic task, they retain the ability to discriminate this difficult contrast at 9 months of age when each individual speech sound is paired with a specific object. For contrasts that completely fall outside any experiences that the infant may have, such as English-learning babies being tested on Cantonese tones, it was not enough to hear the contrast paired with an object. The addition of referential information was required, which involved first showing infants familiar objects paired with their labels (banana, car).32 Moreover, the ability to discriminate tonal contrast was further mediated by vocabulary size. That is, only 9-month-olds with a larger vocabulary were able to discriminate this difficult contrast, even when it was paired with an object. Work by Feldman and colleagues have highlighted the importance of the word in influencing discrimination. They found that 8-month-olds could distinguish similar vowels only after hearing those vowels consistently in distinct word environments.33 Modeling work supports the usefulness of lexical knowledge in determining sound categories. Work by Martin et al. has shown that learners use top-down lexical information for learning phonemes without first needing to learn words by building an initial, proto-lexicon that consists of high-frequency units (n-grams) which are present in the speech signal.34 Together these findings suggest that a number of factors, including distributional, visual, lexical, and vocabulary size and referential cues, contribute to whether infants can discriminate speech sounds. It also leads to the possibility that infants are not directly learning about the speech sounds in their native language, but rather that they attend to information that is related to language more generally, such as the visual, indexical, and linguistic context.

Visual and Sensori-Motor Influences

In many studies cited above, infants demonstrate the ability to discriminate between speech sounds (and nonspeech sounds), in the absence of visual cues and before they have developed the motor abilities necessary for producing speech or the ability to articulate the speech sounds being discriminated. For example, while English-learning infants are able to discriminate the Hindi dental/retroflex contrast between 6 and 8 months of age, at this stage infants have not yet started producing contrasts in words to indicate meaningful differences in a language. Even so, studies have demonstrated that listeners accommodate dynamic and multi-sensory input during speech processing. A classic example of audiovisual integration is the McGurk effect which describes the integration of conflicting audio (ba) and visual information (ga) into a single percept (da). This effect is seen not only with adults35 but also with infants aged 4–5 months.36,37 Therefore it is important to understand how learners represent and integrate multiple cues in processing language, even if these cues are not necessary for speech perception.

Studies looking at visual input have found that learners are sensitive to visual information in discriminating languages as young as 4 months of age.38 Sensitivity to matching auditory and visual cues is seen even earlier39 and recently shown with neonates by Coulon et al.,40 who found that newborns imitate audiovisual congruent facial gestures, but not incongruent audiovisual speech. While evidence for audiovisual integration is found very early, these cues develop throughout infancy and childhood before reaching adult-like processing. Other research has demonstrated that audiovisual integration starts out independent of language background, but narrows with language experience by the end of the first year.41,42 Different neurophysiological measurements are found for matching and mismatching audiovisual information with infants as young as 543 and 10 months of age.44 However, the full maturation of these mechanisms and neural substrates further develops with experience. Research has also examined infants’ attention to faces producing speech at different points in development. Recently, Lewkowicz and Hansen-Tift45 demonstrated that at 8 months infants’ attention shifts from the eyes to the mouth, coinciding with the onset of babbling. While this shift might be driven by concurring speech production skills, the effect is variable across infants.46 The variability observed in infants might in part stem from the diversities seen in early speech production skills which could be addressed in future research using both perceptual and production measures, as reviewed below. This is new emerging literature and the implications of the relationship between production and visual attention will be interesting to explore once there is more available research.

Recent work has also looked at the role of sensori-motor information in speech perception, exploring whether audiovisual perception may involve
intermediate articulatory representations. Work with adults manipulated participants’ articulatory movements during speech perception and found similar effects with production as seen with the McGurk effect. Most recently, Yeung and Werker found related effects with 4 and a half-month-old infants by directly manipulating infants’ sensori-motor information from lip movements. Infants either chewed on a toy, which resulted in lip spreading or they sucked on a pacifier, which resulted in lip rounding. Infants were then presented with audiovisual speech of a woman’s face with either /i/ or /u/. Infants showed a contrasting effect, looking less when their sensori-motor patterns matched the audiovisual display. Neurological evidence to support the link between speech production and perception comes from studies showing patterns of neural activation in the interior frontal region and superior temporal region of infants aged 6 and 12 months.

Other research on the role of sensori-motor factors in speech perception have looked at whether infants’ speech perception abilities correlate to their production abilities, as defined by vocal motor schemes (VMSs), which are frequent consonants in an infant’s production repertoire. DePaolis et al. found that 10-month-old infants with multiple VMSs showed a novelty listening preference for passages containing nonwords comprised of consonants they did not consistently produce. Corresponding results have been reported with infants acquiring Italian and English and Welsh. Such relationships between the domains of perception and production have been captured with an ‘articulatory filter’ where produced sounds are more salient in the learner’s input, which has been argued to play a role in early word learning. Words comprised of produced sounds may be facilitated by previous production experience, which can support the memory and recognition of word forms with those segments. Evidence to support this hypothesis comes from studies of child and adult word learning, which show better word recognition for nonwords that are produced rather than just heard during training. Production might not always be advantageous, as when nonwords are paired with a visual referent, children show better recognition for nonwords that are heard rather than produced during training. This suggests that the integration of multi-sensory information partly depends on the demands of the task. Work has also shown varying sensitivities to structures in perception which parallel children’s production abilities. Research has also begun to examine the role of self-monitoring or auditory feedback in speech, indicating changes from toddlerhood to childhood. We return to the issue of production in our discussion on the development of phonological and lexical representations.

### BOX 1

**BROADER INDICES OF LANGUAGE DEVELOPMENT**

These pivotal findings continue to inspire work in the area of infant speech, and researchers have begun to explore the relationships that might exist between motor, auditory and visual perception, and broader indices of language development, such as vocabulary size. There is evidence that developmental changes in speed of oro-facial movement in infants between 9 and 21 months is correlated with developmental advances in language and cognitive skill. Infants’ ability to discriminate two acoustically distinct vowels (/u/ and /y/) at 6 months is correlated with language abilities at 13–24 months of age. Infants at 5 months who prefer listening to their native-language stress patterns show larger vocabularies at 12 months, and the amount of time 12-month-olds spend listening to speech is related to vocabulary size at 18 months. Other speech processing abilities, such as speech segmentation at 12 months, correlate with expressive vocabulary at 24 months and language skills at 4–6 years. Performance on word–object associative learning with minimally differing words (e.g., bih, dih) at 17 and 20 months is related to children’s performance on standardized tests of language comprehension and production 2 and a half years later. Thus, early multi-sensory skills are tied to later lexical development. These broader indices are further tied to lexical access and processing during spoken word comprehension and spoken word production.

**CURRENT RESEARCH: ASPECTS OF EARLY WORD LEARNING**

Throughout early development infants are demonstrating sensitivity to word forms by showing recognition of familiar and frequent forms and by learning to associate novel words to referents. This section focuses on some of the research demonstrating that infants have emerging lexicons and that within the first 2 years of life, both the recognition and learning of words occurs. Thus, we see evidence that lexical knowledge can support the learning of the sound system and knowledge of the sound system can influence word learning (Figure 1).
Word Recognition

Numerous studies have been conducted which explore infants’ ability to detect mispronunciations of familiar words at various ages. Being able to detect an incorrectly pronounced form provides support for the hypothesis that infants have detailed representations of word forms. These studies have shown that depending on the type and position of the mispronunciation in the word, the age of the child, and the specific task, infants’ representations of familiar words is fairly specific, at least in onset (word-initial) position. For example, Jusczyk and Aslin\textsuperscript{68} found that infants at 7 months do not recognize a familiar word (e.g., cup) if the first segment is mispronounced (e.g., tup). Early studies, using the head-turn preference procedure (HPP), found that sensitivity depends on where in the word and in what type of syllable (stressed, unstressed) the mispronunciation takes place. In this task, infants are typically presented with lists of spoken words either correctly or incorrectly pronounced. Dutch-learning 11-month-old infants prefer correctly pronounced words to mispronounced words where the word-initial segment changed in POA, but when POA is changed for a word-final segment, infants do not show a preference for the correct pronunciation over the mispronunciation, suggesting better specification of consonants in word-initial position.\textsuperscript{69} However, if the onset consonant in an unstressed syllable is changed (e.g., canárd ‘duck’ to ganárd) then 11-month-old French-learning infants treat both words as familiar, that is, they do not notice the mispronunciation.\textsuperscript{70} This is not the case, however, if the phonetic detail occurs in a stressed syllable.\textsuperscript{71} Moreover, by 14 months it appears that infants become sensitive to mispronunciations even in word-final (coda) position.\textsuperscript{69}

Studies using a different task, such as the intermodal preference looking procedure (IPL) (also called the looking while listening paradigm) which presents a visual referent along with the word, have found that sensitivity to a number of mispronunciations is evident by 14 months, even for very similar sounding words. When presented with a display of two known objects, infants of 14–23 months will shift their gaze and look longer toward a target object (e.g., a baby) when they hear its correct pronunciation as opposed to an incorrect pronunciation (e.g., vaby).\textsuperscript{72,73} White and Morgan\textsuperscript{74} found that 19-month-olds demonstrate graded recognition of familiar words depending on the degree of mispronunciation (1–3 phonetic features). When presented with a familiar and novel object, infants tolerated the single feature mispronunciation, but with more featural changes, they no longer looked to the familiar object. Support for detailed representations is also found in 12- and 17-month-old French-learning infants. That is, French-learning infants can distinguish between words differing by only one or two distinctive features of the initial consonant.\textsuperscript{75}

Together, these word recognition studies with familiar words demonstrate that infants’ representations encode at least some phonetic detail, and that degrees of phonological differences (at least in terms of number of features) factor into whether a mispronunciation will be detected. However, this depends on the task, suggesting that a phonological featural system does not necessarily need to be accessed (though some work has argued for early feature representations\textsuperscript{76}).

A new, current line of research has begun to explore how infants deal with atypical speech in terms of accents and/or dialectical differences in word productions. Best et al. describe two ways in which infants must manage variable input.\textsuperscript{77} For phonological distinctiveness, infants must determine when a sound difference signals a change in word meaning (e.g., beet, bit). On the other hand, they
must also be able to recognize a word even if it has an atypical pronunciation (e.g., pen pronounced as pin), which is known as phonological constancy. Indeed, by 19 months, infants demonstrate phonological constancy by preferring accented pronunciations of familiar toddler words to words more familiar to adults.\textsuperscript{77} Interestingly, when just listening to speech in familiar or novel accents, 6-month-olds prefer to listen to the familiar accent while 9-month-olds do not show any preference.\textsuperscript{78} This suggests that even without specific word knowledge, by 9 months, infants accommodate for accented speech. More recent work has also suggested that young children’s emerging abilities to accommodate accents can vary depending on the task,\textsuperscript{79–81} but as children’s vocabulary develops their abilities become more flexible and more proficient.\textsuperscript{82,83} This ability to accommodate for variation in the input could suggest that stable phonological representations of word forms are emerging or that knowledge of allophonic variation (different realizations of speech sounds) allows children to ignore a certain degree of variation in how a word is produced (see Seidl and Christia’s recent review\textsuperscript{84}).

**Learning New Words**

Phonological knowledge appears to guide infants’ word–object associative learning. Around 12 months of age, infants demonstrate knowledge of what appears to be a good object label based on the phonological structure of the word. This learned bias for object labels to conform to noun-like phonological patterns is observed when infants are not provided with any contextual or referential cues. In a study using the ‘switch’ procedure,\textsuperscript{85} infants observed a novel object moving back and forth on a screen while hearing a repeated novel label until they habituated. Infants were then presented with two test trials, a ‘same’ correct word–object pairing, and a ‘switch’ incorrect pairing trial. Using this task, English-learning 12-month-olds accept well-formed object labels, such as fep and wug, but reject isolated speech sounds (e.g., /f/) and communicative sounds (e.g., oooh and shhh) as possible labels.\textsuperscript{86} That is, infants recover, or look longer to the ‘switch’ trial in the case of well-formed object labels, but not in the case of the others. Moreover, infants of this age also prefer to map these noun-like labels to objects more so than function-like phonological patterns, such as iv and keh.\textsuperscript{87} Together these findings suggest that even though English-learning infants have heard various speech sounds, communicative sounds, and even function words, they have acquired lexical knowledge about what constitutes an appropriate phonological pattern for content-like, and more specifically, noun-like object labels. Interestingly, a developmental narrowing appears to occur between 14- and 20 months of age in infants’ awareness of what an appropriate label may be in their native language and in infants’ willingness to accept forms that contain illegal sounds as labels for objects. May and Werker\textsuperscript{88} found that when infants were provided with a referential training phase, 14-month-olds, but not 20-month-olds, will map a word form containing a nonnative speech sound, such as a click, to an object, suggesting that when young infants are provided with an additional cue that clarifies the nature of the switch task, 14-month-olds are still willing to map a label that contains an illegal phoneme to a novel object. The importance of word forms is demonstrated in a study using the anticipatory eye-movement paradigm to evaluate how learners weigh word forms and referents in early word–object associations. Participants were trained on novel word–object associations that were decoupled at test to create ambiguous stimuli which pitted word forms and objects against each other. Results indicated that at 18 months, infants preferentially weigh word forms over objects in early word–object associations, where the reverse pattern was found with adults.\textsuperscript{89}

Infants of 12 months have also begun to learn what is considered a legal phonotactic sequence for their language’s object labels. English-learning infants will not accept illegal onset sequences, such as the Czech form ptak, as a possible label.\textsuperscript{90} However, they are still somewhat flexible in accepting atypically sounding forms, such as the Japanese form sika, which contains legal sound sequences for English, but is phonetically atypical for English. Work with 18-month-old infants has shown similar findings.\textsuperscript{91} This suggests that perhaps the variability experienced in the ambient language input allows for a certain amount of phonetic variation to be acceptable, but going too far afield in terms of the legality of sound sequences is not acceptable.

While infants of 12 months have developed a strong sense of what constitutes a reasonable phonological form for an object label, they are not always able to use their knowledge of speech sound categories to learn novel word–object pairings that differ only slightly from one another in the switch task. Even at 14 months of age, infants have difficulty in learning labels, such as bin and din, when paired with two novel objects.\textsuperscript{66,92} This suggests that when trying to learn, map, and remember a new pairing, if the word forms are too similar, then infants have difficulty with the task. Thus, even though novice word learners have no difficulty in discriminating /b/ and /d/ with the limited resources available to them something has to
Difficulty in detecting minimal pair contrasts is not limited to consonants, although very few studies have specifically compared infants’ use of vowel versus consonant information in early words. Using an inter-modal preferential looking task, Mani and Plunkett found that 18-month-olds recognize mispronunciations of word-medial vowels in familiar words, but this finding is less robust at 15 months, whereas infants of both ages notice consonant mispronunciations. This pattern of results may indicate that vowels are not used as efficiently as consonants in early word learning. Indeed, when 15-month-olds are taught novel words differing only in their vowels, they detect the ‘switch’ if the vowels differ in the height dimension (e.g., /i/ and /I/), but not if they differ primarily in backness (e.g., /i/ and /u/), suggesting that some featural differences are easier to detect than others.

Similarly, results from an interactive object categorization task that uses novel words show that infants as old as 20 months still might confuse labels that minimally differ in their vowel, but succeed in distinguishing labels that minimally differ in consonants, suggesting different processing of vowels and consonants. However, 18-month-olds are sensitive to small changes in vowels in familiar and novel words, and infants of 14 months can notice broad vowel mispronunciations (i.e., three feature changes) in newly learned words. Thus, the picture relating to novice word learners’ use of vowel information is not as clear as with consonants. This could be due to a number of factors, such as the use of different tasks (categorization, habituation), vowel contrasts (including multiple-feature contrasts), and word familiarity (familiar vs novel).

The range of findings in the word–object associative learning literature suggests that infants are actively engaged in learning words, but that sensitivity to various phonological properties of words depends on the task. Indeed, the relationship between the emerging lexicon and sound system is dynamic and evolves throughout development. Thus, while infants may be actively engaged in learning the phonological system of their language, phonology also emerges from the experience of learning about words.

In this last section we turn to early phonological and lexical influences in early language development, keeping in mind their bidirectional relationship.

### Phonological and Lexical Representations

Examinations of phonological and lexical relationships are relatively new. Adding to the complexity of understanding these relationships, various methodologies result in different findings. It has recently been argued that there is a bidirectional or reciprocal relationship between children’s developing phonological and lexical knowledge. For example, in the Processing Rich Information from Multidimensional Interactive Representations (PRIMIR) framework, phonemes emerge from generalizations made across dense, meaningful words or lexical neighbors (also see work on the Lexical Restructuring Hypothesis). In turn, the emergence of phoneme representations influences how learners process speech and learn new words. We review evidence that provides insights into learners’ early phonological and lexical representations, specifically looking at phonotactics and neighborhood densities, though some studies have considered other cues such as word frequency. At the end of this section, we briefly discuss the effect that children’s own productions may have on the development of phonological and lexical representations.

### Phonotactics

As mentioned above, the development of phonotactic knowledge begins around 9 months of age and the earliest studies measured phonotactic knowledge with preferential looking paradigms. It was discovered that this knowledge can help learners discover other aspects of language such as locating word boundaries and classifying lexical items into different word classes. Other research with older children has investigated whether a word’s phonological pattern impacts how it is learned. Findings show that 18-month-olds are able to learn nonwords with legal phonotactic patterns, but do not learn items which violate English phonotactics. More recently, MacKenzie et al. demonstrated similar phonotactic effects in the mapping of word–object associations with 12-month-old infants. Children’s productions also obey language-specific phonotactic patterns. Neurophysiological measures (N400) are reported at 19 months for phonotactically legal nonwords, but not for phonotactically illegal nonwords.

Other studies have manipulated phonotactic frequency or PPs rather than phonotactically legality. Gonzalez-Gomez et al. found that 14-month-old French-learning infants learned nonwords with frequent phonotactic patterns more easily than nonwords with less frequent structures. By 16 months, infants were able to learn both types of nonwords. In other studies using the preferential looking task, significant differences were reported in 2-year-olds’ ability to detect mispronunciations on high PP (but not low PP) words and nonwords. Probability-based effects in perception are also reported in 2-year-old
Dutch-learning children who demonstrate the ability to detect segmental contrasts in both high and low PP nonwords; however, contrasts in high PP nonwords were perceived more accurately.\textsuperscript{115} Words with high-frequency phonotactic patterns also show an advantage in word learning,\textsuperscript{116,117} although this effect changes to an advantage for low PP forms by 5 years\textsuperscript{118} and with adults.\textsuperscript{119} When looking at young children’s production accuracy, many studies have reported that children are more accurate at producing frequent structures compared with less frequent structures\textsuperscript{120–123} and produce them with shorter naming latencies.\textsuperscript{124} Together, these studies illustrate that words’ sound properties influence speech processing and lexical acquisition: early in development, legal and frequent sound patterns are at an advantage compared with illegal or less frequent sound structures.

### Neighborhood Densities

Studies have also revealed that learners are not only sensitive to phonological factors, but also to lexical factors, as measured by phonological neighborhood density (ND). This refers to the degree of sound overlap a word has with other words. The initial research on this topic was to investigate the level of segmental detail in children’s lexical representations. It was argued that children’s early representations are holistic and undergo restructuring to accommodate segmental overlap across words.\textsuperscript{105} While research has now shown that children’s early lexical representations encode phonetic detail, the more recent PRIMIR model still holds that similarities across words in the lexicon lead to the emergence of phonemes.\textsuperscript{4,101} Thus, one area of research to explore the bidirectional relationship between phonological and lexical development is to examine how words may cluster in children’s early vocabularies.

Researchers have also examined NDs to investigate the learnability of words depending on their degree of overlap with other words in the lexicon. A high degree of sound overlap can be both advantageous and disadvantageous. It may be easier to maintain similar sound words in working memory, and at the same time, it may be difficult to perceptually distinguish between similar sounding words.\textsuperscript{123,126} Evidence for advantageous effects have been documented. Storkel\textsuperscript{127} found that children between 16 and 30 months are more likely to have acquired shorter words with high ND compared with longer words with low ND, although children show variability in this pattern, suggesting that the effects of ND vary at different points in development.\textsuperscript{128} In other work based on analyses of English, European French, and Danish, researchers found that children with smaller vocabularies have more high ND words which decreased as vocabulary size increased.\textsuperscript{106,107,129} Words from dense neighborhoods are produced more accurately and with less variability compared with words having low ND.\textsuperscript{130,131} However, McKean et al.\textsuperscript{118} found the reverse pattern of an advantage for low ND in word learning with children starting at age 3 (also see Swingley and Aslin\textsuperscript{132}). While results from studies of ND may appear at times to be contradictory, the effects vary with how one measures children’s knowledge. The impact of ND on the child’s responses depends on whether researchers have looked at perception/comprehension, word learning, production influences and whether data are based on laboratory studies or naturalistic data.

The above studies report on specific effects of phonotactics and neighborhood densities, and most studies have not differentiated between PP and ND. As such, in many studies the stimuli have been confounded. While it has been generally argued that studies with nonwords show PP effects and studies with real words show effects of ND, some recent work has looked at the separate effect of these factors and how this changes during development, as children’s lexical and phonological knowledge grows. Using the PRIMIR model, McKean et al.\textsuperscript{118} created nonce stimuli that varied orthogonally on PP and ND, and predicted that early in development, words with high ND or large sound overlap (pin vs bin) will show inhibitory effects because the learner cannot rely on more efficient phoneme level representations to discriminate the words. At this stage, the task is very demanding because the critical dimension or features that define the words are not known to the learner. However, if phonemic representations emerge first from generalizations made across dense, meaningful words, then speech processing should be more robust in these high ND contexts. Indeed, they found that with children between the ages of 3–5 years, the effect of PP changes over development from an advantage for high PP stimuli to low PP stimuli. With ND, there was an overall advantage for low ND. McKean et al. argue that this reflects a functional reorganization in the developing lexicon.

More recent studies have begun to look at other aspects of sound similarities across words, beyond the general and/or global measure of ND. Some research work has looked at the position of sound overlap in English, French, and Dutch and shown that in children’s early lexicons, many words contrast in word-initial position compared to other word positions, similar to distributions in the ambient language.\textsuperscript{133–135} Studies have also examined
phonological priming effects and found that the direction of sound priming in spoken word comprehension changes between the ages of 18 and 24 months, shifting from facilitative to inhibitory priming.\textsuperscript{136,137} The authors argue that the facilitatory effect at 18 months reflects the organization of lexical representations: words are not yet stored in dense cohorts leading to the absence of lexical interference. By 24 months, these phonological cohorts have emerged, leading to lexical interference effects from more similar sounding words. This leads to the disruption of word recognition in dense neighborhoods compared to sparse ones. Other research combining multiple measures of sound relationships in the lexicon, show that the likelihood that children will produce a word not only depends on local measures of phonological properties (traditional definition of ND), but it also depends on how a word fits into the rest of the lexicon. Together, these properties are argued to support lexical access for speech production, an area which to date has not received much attention in the literature.\textsuperscript{126}

The second year of life marks a transition stage, when learners move from comprehension based knowledge in infancy, to a system that allows for both the comprehension and production of language. The emergence of spoken word production is amazing, considering the fact that young children are not born with the ability to talk, and yet they start producing words within their first 12 months. Above we highlighted the bidirectional nature of phonological and lexical development. However, we must also consider how children’s emerging speech production skills may impact the development of phonological and lexical representations. Exemplar theories and usage-based models propose that speech output representations are integrated into phonological and lexical representations.\textsuperscript{102,131,138,139} In these models, the development of sound, word representations, and speech is multidirectional as speech output is a part of long-term representations. Developing new methods to examine these relationships will be a challenge for the field because it will require online measures or more implicit measures of speech output with children who have just started producing language.\textsuperscript{140}

**CONCLUSION**

Infants’ sensitivity to various aspects of the sound system develops over the course of the first year of life. As they begin to segment, recognize, and learn words, they begin to use the phonological patterns observed in the input and patterns across words stored in their developing lexicons to subsequently process new information. This tight relationship between infants’ lexicons and their sound system is useful for constraining the search space for possible word forms. Indeed this makes sense as words and phonology develop contemporaneously. Recent research in support of this view comes from studies looking at learner’s early ‘proto-lexicons’.\textsuperscript{34} Rather than adopting a strictly bottom-up view of learning, this research explores additional top-down ways in which phonological regularities in a language are learned across proto-lexicons based on frequently occurring sequences in the language. This approach predicts that learner’s proto-lexicons will include sequences that correspond to both real words and nonword sequences in the language. Evidence to support this hypothesis comes from experiments demonstrating that 11-month-old infants prefer high-frequency nonword sequences to low frequency nonword sequences and do not distinguish between high-frequency words and high-frequency nonwords.\textsuperscript{141}

We leave off with an interesting observation. To date, research work has primarily looked at the overall size of the learner’s lexicon in terms of looking for connections between phonology, vocabulary, and age. However, if phonological categories also emerge from the lexicon, the overall size of a child’s lexicon might not be as important as the sound properties of the words in a child’s lexicon. Future studies examining the structure and contents of early developing lexicons will help to further elucidate the relationship between the emerging sound system and lexical development.

**REFERENCES**


© 2014 John Wiley & Sons, Ltd.
76. van der Feest SVH, Fikkert P. Building a phonological lexicon. (Manuscript under review).
102. Munson B, Beckman ME, Edwards J. Phonological representations in language acquisition: climbing the ladder of abstraction. In: Cohn A, Fougeron C,


112. Gonzalez-Gomez N, Poltrock S, Nazzi T. A “Bat” is easier to learn than a “Tab”: effects of relative phonotactic frequency on infant word learning. PLoS One 2013, 8:e59601.


**FURTHER READING**


