Review in Advance first posted online on October 16, 2015. (Changes may still occur before final publication online and in print.)

Constructing a Proto-Lexicon: An Integrative View of Infant Language Development

Elizabeth K. Johnson

Department of Psychology, University of Toronto, Mississauga, Ontario L5L 1C6, Canada; email: elizabeth.johnson@utoronto.ca

Annu. Rev. Linguist. 2016. 2:18.1–18.22

The *Annual Review of Linguistics* is online at linguist.annualreviews.org

This article's doi:
10.1146/annurev-linguistics-011415-040616

Keywords

infant speech perception, word learning, word recognition, language acquisition, word segmentation problem, phonological development

Abstract

Infants begin learning the phonological structure of their native language remarkably early and use this information to extract word-sized chunks from the speech signal. While acquiring the language-specific segmentation strategies appropriate for their native language, infants are simultaneously beginning to form word–object pairings and learning which sound contrasts are meaningful in the native language. They are also working out how to assign words to word classes, paying attention to the use and placement of function words, and learning how speakers of the language string words together to form sensible grammatical utterances. Amazingly, infants tackle all of these tasks simultaneously, with success in each of these domains dependent on success in the others. This review focuses on infants' discovery of word forms in speech, their construction of a proto-lexicon, and the development of linguistic knowledge during their first year and a half of life. By discussing the development of lexical knowledge in relation to other aspects of linguistic development, I demonstrate the advantages of an integrative approach to understanding early language acquisition.

1. INTRODUCTION

Infants typically utter their first word shortly before their first birthday, marking an important developmental milestone in childhood. But it would be inaccurate to think of the preverbal infant as prelinguistic. In the first 12 months of life, infants are quietly extracting linguistically relevant regularities from the speech signal. By 1 year of age, they have already silently passed many important language-learning milestones, including acquiring the sound structure of their native language(s), attaching meaning to a small cohort of frequent word forms, and gathering the rudimentary knowledge needed to understand how words combine with other words to form sentences. Around their second birthday, children typically know more than 300 words (Fenson et al. 1994) and are exhibiting increasingly efficient performance in online comprehension tasks (Fernald et al. 2006). Children are also using grammatical knowledge to learn new words (e.g., Gerken 2002), coping effectively with unfamiliar accents (e.g., Mulak et al. 2013, van Heugten & Johnson 2014) and starting to produce multiword utterances (e.g., Brown 1973). In short, children begin acquiring their native language far earlier than their overt behavior suggests, and they do so incredibly efficiently.

In the past 30 years, improvements in infant testing methodologies have enabled researchers to uncover surprisingly sophisticated language abilities in young infants (e.g., Aslin et al. 2015, Fernald et al. 2008, Johnson & Zamuner 2010). In this review, I draw on some recent discoveries in this area to address one of the most important questions in the field: How do infants transition from hearing speech as a string of meaningless sounds to perceiving speech as a string of recognizable words? And how does the acquisition of word forms relate to other aspects of language development?

2. THE BEGINNING STATE

From the moment they are born, infants are attuned to language. Neonates' brain responses to linguistic stimuli are already lateralized to the left hemisphere (Shultz et al. 2014), and newborns prefer to listen to natural speech than to temporally reversed speech (Peña et al. 2003). They also prefer infant-directed speech to adult-directed speech (Cooper & Aslin 1990) and singing voices to musical instruments (Cairns & Butterfield 1975). As infants' exposure to their native language builds up, they benefit from built-in listening biases and powerful learning mechanisms that help them focus on those regularities that are most meaningful in the native language.

Understanding what aspects of linguistic knowledge are innate and what aspects are learned is a classic question in the field of language development (e.g., Johnson 2012, Lidz & Gagliardi 2015, Yang 2004). But even if one were to assume no innate linguistic knowledge in humans, the language-learning newborn still would not be a blank slate. The fetal auditory system begins functioning during the third trimester of pregnancy, allowing some environmental sounds (including a low-pass filtered version of the mother's voice) to pass through the mother's body to the womb (e.g., Lecanuet & Schaal 2002). This allows the human fetus to get a jumpstart on learning her native language by eavesdropping on her mother in the months preceding birth (Saffran et al. 2006). The low-pass filtered speech to which the fetus is exposed carries information about the rhythm and intonation of language and perhaps some vowel information. Remarkably, human fetuses appear to retain memories of the language exposure they receive in the womb. Rhymes and songs heard in the third trimester are recognized after birth (DeCasper & Spence 1986, Partanen et al. 2013), and newborns also recognize the rhythm of their native language (e.g., English-learning babies prefer stress-timed English over syllable-timed Spanish, whereas Spanish-learning babies show the opposite preference; Moon et al. 1993, Nazzi et al. 1998). Additional evidence of early prosodic knowledge originating from prenatal experience is provided by crosslinguistic investigations of newborns' cries. French newborns cry in a rising melody, argued to be reflective of the French language, whereas German newborns cry in a falling melody, argued to be reflective of the German language (Mampe et al. 2009). In short, newborns arrive in the world with a sturdy foundation for language development already in place.

Although the low-pass filtered speech that reaches the womb allows the human fetus to learn a great deal about the prosody of her native language, this speech does not carry much segmental detail. Nonetheless, infants' inborn perceptual capabilities are set up to enable rapid acquisition of this information soon after birth. For example, young infants possess categorical discrimination for stop consonants (Eimas et al. 1971) and a universal sensitivity to the acoustic cues distinguishing most of the contrasts used in the world's languages (Trehub 1976, Werker & Tees 1984; see also Aslin et al. 2002, Jusczyk 1997). The contrasts that infants initially struggle to perceive tend to be acoustically subtle and less common across the world's languages (e.g., Burnham 1986, Narayan et al. 2010).

As infants gain more experience with language, they transition from universal to language-specific listeners (Werker & Tees 1984). By age 2 to 3 months, infants have begun imitating adults' productions of vowels (Kuhl & Meltzoff 1996) and can detect the link between segmental information presented in the visual and auditory streams (e.g., Patterson & Werker 2003). Although there is very limited evidence of infants' attunement to the native language inventory at this age, cross-cultural adoption studies have shown that language exposure in these early months permanently alters how listeners process speech (Choi 2014; see also Singh et al. 2011). By age 6 months (and perhaps even earlier; see Moon et al. 2013), infants have begun to exhibit language-specific vowel perception (Kuhl et al. 1992, Polka & Werker 1994). Attunement to native language consonants takes a little longer (see Cutler & Mehler 1993 for a related discussion). At around 10 to 12 months, infants generally show a heightened sensitivity to phonetic contrasts that signal meaning differences in the native language, along with a decline in sensitivity to phonetic contrasts that do not signal a meaningful difference (Werker & Tees 1984; but see Tyler et al. 2014 for a more nuanced view of this process). Factors such as acoustic salience and frequency appear to influence how quickly this process occurs for specific contrasts (e.g., Anderson et al. 2003, Burnham 1986, Narayan et al. 2010). During this period, infants are also tuning into other important properties of their native language, such as whether tone (Mattock et al. 2008) and lexical stress (Skoruppa et al. 2009) are used contrastively.

How can infants learn so much about their native language in such a short time? Even with prenatal exposure to language in the womb and inborn constraints, the speed with which children tune in to the phonological structure of their native language is certainly impressive. And in the latter half of the first year of life, infants use this phonological knowledge to construct a proto-lexicon and begin to learn about the ordering of words in the native language. As I discuss further in the following sections, a secret to infants' success in acquiring language may be their integration of information across distinct domains of language knowledge (e.g., the use of lexical information to work out the native language phonology and syntax, and vice versa).

3. WHEN ARE FIRST WORDS LEARNED?

If one were to go to a local playground and ask half a dozen parents when their children learned their first word, one would get a wide variety of responses. Some parents would boast that their children are language-learning prodigies who said 'mom' at 3 months of age. Other parents would report that their children were virtually mute until well after their second birthday. Some variation in parents' responses can certainly be attributed to individual variation in the onset of word production by different children (Fenson et al. 1994, Labov & Labov 1978), but much of it is also due to parents' different notions of what it means to "learn" a word (Styles & Plunkett 2009, Tomasello & Mervis 1994).

Language researchers also have different criteria for defining what it means to learn a word (e.g., Vihman & McCune 1994). Researchers interested in production focus on when words are first spoken, whereas researchers interested in perception focus on when the sound pattern of a word is first recognized (regardless of whether it is understood or not). Still other researchers have even stricter definitions for when a word is learned, requiring words to be used flexibly in different contexts or specified in terms of abstract phonemes. In truth, all of these definitions for when a word is "learned" are legitimate because word learning is not an all-or-nothing affair. Children often recognize word forms as familiar before they attach a meaning to the word, and children's phonological and semantic representation of words changes over the course of development. For example, children may recognize the general sound pattern of a word from repeated exposure, and know the word form is likely a noun on the basis of its sentence placement (e.g., Höhle et al. 2004, Shi & Melançon 2010), but they may still not know the precise meaning or phonological structure of the word for many months (see Swingley 2009 for a review).

If we were to set the bar as low as possible for what it means to learn a word, we would find that infants show evidence of having learned their first word by 4.5 months of age. That is, by age 4.5 months infants preferentially listen longer to repetitions of their own name than to repetitions of another infant's name (Mandel et al. 1995). However, there is no indication that very young infants know what their name means, or that they possess fully specified representations of the sound patterns of their name (Bouchon et al. 2014).

If we were to set the bar slightly higher and require that infants have at least some notion of what a word means, the age at which children learn their first word would still be quite early. Eye-tracking studies have shown that 6-month-old infants look to an image of their mother when they hear 'mommy' and to an image of their father when they hear 'daddy' (Tincoff & Jusczyk 1999). They also show some recognition for other frequent words, such as 'hand' and 'foot' (Bergelson & Swingley 2012, Tincoff & Jusczyk 2012). At approximately the same age, children first begin showing evidence of learning word forms and word–object pairings in the lab (Bortfeld et al. 2005, Friedrich & Friederici 2011, Gogate 2010, Johnson et al. 2014, Kooijman et al. 2013, Shukla et al. 2011). Interestingly, infants' early proto-lexicons appear to be overspecified. For example, 6-month-olds look only to an image of their own mother (not another infant's mother) when they hear 'mommy,' suggesting that they do not realize that the term can apply to anyone other than their own mother (Tincoff & Jusczyk 1999). They also fail to recognize newly learned word forms that are altered phonologically (Johnson et al. 2014, Jusczyk & Aslin 1995) or produced in a different voice (Houston & Jusczyk 2000; see, however, Johnson et al. 2014, van Heugten & Johnson 2012), a different emotional affect (e.g., Singh et al. 2004), or a different accent (e.g., Schmale et al. 2010). Understanding how infants overcome this apparent overspecification of items in their proto-lexicon is currently an active area of study.

In the second half of the first year of life, infants' word knowledge continues to mature rapidly. By the age of 8 months, infants store novel word forms in memory for at least 2 weeks (Jusczyk & Hohne 1997), and soon thereafter, many parents report that their children are beginning to understand words (Fenson et al. 1994). By age 11 to 12 months, infants show substantial improvement in their ability to recognize word forms across changes in voice (Houston & Jusczyk 2000) and affect (Singh et al. 2004). Around the same time, children start saying their first words (Fenson et al. 1994), and these early words build upon the production skills acquired in the babbling period (e.g., Keren-Portnoy et al. 2009). Note, however, that there is substantial individual variation in not only when but also how children begin speaking. For example, some children seem more focused on producing whole phrases, whereas others are more focused on individual words (e.g., Peters 1977).

In the months that follow the first birthday, children gradually continue adding words to their vocabulary. Occasionally, the semantic scope of these early words does not neatly map onto the adult form (e.g., initially overextending the meaning of 'dog' to all four-legged creatures; e.g., Bowerman 1976, Rescorla 1980). At approximately 18 months, the rate at which toddlers add words to their vocabulary tends to accelerate, a phenomenon termed the vocabulary spurt. Some researchers (e.g., Nazzi & Bertoncini 2003) have argued that the vocabulary spurt marks the emergence of "real" words that are semantically and phonologically mature—that is, that word learning that occurs before the vocabulary spurt is qualitatively different from the learning that occurs after it. Others (e.g., McMurray 2007) have disagreed.

In summary, acquiring a word does not happen in an instant. Rather, word learning is better understood as a gradual process, with different dimensions of children's lexical representations being updated and refined over time. A child may recognize the word form 'dog' by 6 months, and understand that it belongs to a category of words typically preceded by determiners by 14 months, but she may not fully understand until many months later that both Fido and Spot are called 'dogs' but the cat next door is not. The fact that children add words to their vocabulary in a gradual fashion makes it difficult to define when precisely a word is "learned"; however, it provides a wealth of clues regarding how children acquire language.

4. HOW ARE FIRST WORDS LEARNED?

Perhaps the most obvious challenge facing the word-learning child is working out what a word means. When mom utters 'Pinocchio,' what does this word (or phrase) refer to? Is mom referring to the puppet she is holding, or the jumping motion she is making with the puppet, or the fact that the puppet has three fingers instead of five? Or is mom simply asking what we would like for lunch this afternoon? Without some strategies for working out what the most likely referent is for a word form, the child is faced with a virtually infinite number of possible mappings. There is an enormous literature on this topic, known as the gavagai problem (see Bloom 2001 for a review). But before (or perhaps while) children are solving the gavagai problem, they must extract word-sized units from speech. The task of extracting word forms from speech is fully as complicated as working out what a label refers to. In this section, I discuss both of these processes, then present some questions regarding the relationship between word–referent pairings and so-called real words.

4.1. Learning Word Forms

When adults hear speech, words seem to naturally pop out as discrete entities, like beads on a string. But silences between words, analogous to the white spaces between words on this page, do not exist—spoken words run into each other, blurring word boundaries (Aslin et al. 1996, Cole & Jakimik 1980). The illusion of physical word boundaries in our native language is caused by our knowledge of what words typically sound like (see Cutler 2012 for a review). This is why it is nearly impossible to identify where one word ends and the next begins when listening to an unfamiliar language. So if adults hear words in their native language only because they already know what words sound like, what cues do children use to first find words in speech?

Finding words in speech despite the lack of reliable acoustic cues to word boundaries has been termed the word segmentation problem. Adults use their knowledge of how words sound in the native language as a heuristic to solve this problem. For example, adult English speakers are biased toward perceiving strong syllables as word onsets because most content words begin with a strong syllable in English (e.g., Cutler & Norris 1988). Adults also use probabilistic phonotactic knowledge to identify likely word boundaries (e.g., McQueen 1998). That is, they use their knowledge of which phonemes occur in certain positions within and across word boundaries to detect word boundaries (e.g., in English, the sequence /mt/ occurs much more frequently across word boundaries than within words). In addition to using language-specific phonological knowledge, adults may use frequently occurring function words to segment speech (e.g., if I know 'the' is a word and I hear 'the ball,' then I can infer that 'ball' is a word; Christophe et al. 1997).

Infants begin using many of these same language-specific word segmentation strategies early in development. For example, by age 7.5 months, English-learning infants are readily segmenting strong–weak words (e.g., 'kingdom' or 'hamlet') but not weak–strong words (e.g., 'device' or 'guitar') from speech (Johnson & Jusczyk 2001, Jusczyk et al. 1999). By age 9 months, English learners are using phonotactic cues to find word boundaries (Mattys & Jusczyk 2001). And shortly before their first birthday, English learners are using function words to locate words in speech (Kim & Sundara 2014, Shi et al. 2006). Infants learning other languages show similar patterns in the acquisition of language-specific segmentation strategies (e.g., Houston et al. 2000). How do infants come to learn these language-specific segmentation strategies? Clearly, this knowledge cannot be inborn, as words pattern differently in each human language (e.g., Polish words are stressed on the penultimate rather than the first syllable; see Peters 1981 for a related discussion).

4.1.1. Do infants really have to solve the word segmentation problem? Perhaps the most obvious explanation for how children find words in speech is that parents solve the segmentation task for them. That is, parents might address their children with predominantly one-word utterances, eliminating the need for infants to have some clever bootstrapping strategy to extract their first set of words from speech. Several corpus studies have been carried out to investigate this possibility. In a study where all conversations directed to (or heard by) a Dutch infant between the ages of 6 and 9 months were recorded, only 7% of the utterances directed to the infant consisted of isolated words (van de Weijer 1998; see also Johnson et al. 2014, Swingley 2005). In another study in which American mothers were brought to the lab and explicitly asked to teach their English-learning 12-month-olds new words, targets were produced in isolation on average approximately 20% of the time (Aslin et al. 1996; see also Johnson et al. 2013). And some word types (e.g., the function words 'a' and 'the') were never produced in isolation. Moreover, mothers varied widely in how often they used one-word utterances (some mothers never produced any isolated words at all).

Clearly, caregivers do not solve the word segmentation problem for infants by speaking almost entirely in one-word utterances. But do parents produce enough isolated words to help infants solve the word segmentation problem? Some researchers have argued that although infants are not addressed predominantly with isolated words, they still hear enough isolated words to support the development of word segmentation strategies (e.g., Johnson & Jusczyk 2001, Lew-Williams et al. 2011; see also Altvater-Mackensen & Mani 2013). Proponents of this view suggest that infants analyze the sound structure of the isolated words in their input and use this information to find more words in fluent speech (e.g., an English-learning infant might notice that most of the words she hears in isolation begin with a strong syllable, and therefore develop a bias toward perceiving strong syllables as word onsets).

Support for this view has been provided by studies showing that those words that mothers produce in isolation are more likely to appear in children's early productive vocabularies (Brent & Siskind 2001). Artificial language experiments have provided additional support for this hypothesis (Lew-Williams et al. 2011). However, a weakness of this proposal is that infants have no way of determining when they have heard a word in isolation (e.g., How does the child know whether 'Pinocchio' is one word, three words, or more?). Thus, attention to isolated words does not seem like the ideal solution to the word segmentation problem. However, see Section 4.1.4, below, for a discussion of a closely related word-finding strategy using utterance edges to learn language-specific word segmentation cues; given that isolated words are simply words flanked by two utterance boundaries, the two strategies are similar.

4.1.2. Using distribution cues to find word boundaries. Another word-finding strategy infants might use has been termed distributional learning. This class of strategies involves tracking the statistical distribution of linguistic elements in the speech stream and using this information to identify likely word boundaries. All of these strategies are based on the notion that words can be defined as statistically coherent sequences of sounds. But these strategies differ in the type of element being tracked and the types of computations being performed over these elements.

Harris (1955) described a phoneme-based distributional learning approach that could help field linguists find morphemes in an unfamiliar language. By tracking how many possible segments could follow any other given segment in the language, linguists could identify likely word boundaries. Similar approaches have been implemented in computational models of infant speech perception (e.g., Batchelder 2002). However, these models appear to be psychologically implausible because they assume that young infants perceive the speech signal as a string of abstract segments that map cleanly onto adult phoneme categories (see Johnson 2012, Jusczyk 1997, Peters 1981, and Rytting et al. 2010 for related discussions).

Others have proposed that infants learn to segment words from speech by tracking the distribution of utterances, not phonemes, in the input. According to this proposal, infants store all heard utterances as possible words, and then use a subtraction method to eventually break down these stored utterances into word-sized chunks (see Brent & Cartwright 1996 for the implementation of this strategy in a computational model). To illustrate this strategy, imagine a child hears 'Look. Look here. Here is the cat.' In this case, 'look' would be postulated as a possible word because it occurs in isolation. Therefore, 'look' would be subtracted from the utterance 'look here,' leaving the possible word 'here.' Then, upon hearing 'here is the cat,' the child would subtract the word 'here' and store the string 'is the cat' in memory. Eventually, new utterances containing the words 'is' and 'the' in different contexts would be heard, allowing the child to find the possible word 'cat.' Behavioral studies with both adults and infants have provided some support for this proposal. For example, adults use this method to find words in an artificial language (Dahan & Brent 1999), and infants use their own names to break up longer utterances (e.g., to extract 'cup' from 'Here is Joey's cup'; Bortfeld et al. 2005, Mersad & Nazzi 2012). However, storing all heard utterances (not only names) in memory would be computationally demanding, and it is not yet clear how effective this word-finding strategy would be for infants.

A final, and perhaps the best known, distributional strategy for finding words in speech has been termed statistical learning. By tracking the baseline frequency of each syllable in the input, as well as how often each syllable is followed (or preceded) by every other syllable, infants could calculate transitional probabilities between syllables [probability of $Y \mid X = $ (frequency of XY)/(frequency of X); Saffran et al. 1996]. Because words can be defined as sequences of syllables that consistently co-occur, dips in transitional probabilities are cues to likely word boundaries. To put it more concretely, imagine the child hears the phrase 'hello baby.' In all of the input heard by the child in the first 6 months of life, the transitional probability between the syllables within 'hello' and 'baby' are likely to be much higher than the transitional probabilities between the syllables 'llo' and 'ba.' Thus, the infant can infer that 'hello' and 'baby' are likely words, whereas 'lloba' is not.

Support for the transitional probability hypothesis has been provided by artificial language–learning studies. In a seminal study (Saffran et al. 1996), 8-month-olds were presented with a synthesized continuous stream of speech containing four repeating CVCVCV trisyllabic words (e.g., 'golatudaropipabikudaropi…'). Between-word syllable transitions were lower than within-word syllable transitions. The prosody of the language was flat, and there were no pauses between words (an impossible feat for a speaker to accomplish, as no human can speak continuously without ever pausing to take a breath). After a brief 2-min exposure to the artificial language, infants could distinguish words from nonwords (i.e., they recognized the sequence 'golatu' as more familiar than the cross-word sequence 'tudaro'). Subsequent studies have reported similar results with 5-month-olds (Johnson & Tyler 2010, Thiessen & Erickson 2013), an age at which infants are not yet sensitive to language-specific cues to word boundaries (Thiessen & Saffran 2003).

For nearly 20 years, tracking transitional probabilities between syllables has been the dominant explanation for how infants first extract words from speech and bootstrap the sound structure of their native language. Numerous artificial language–learning studies have replicated and extended the original findings that infants can extract word boundaries by tracking transitional probabilities (see Romberg & Saffran 2010 for a review) and that language-specific segmentation strategies can then be inferred as a result (Sahni et al. 2010, Thiessen & Erickson 2013, Thiessen & Saffran 2007). A growing controversy in the field, however, has been whether infants' ability to track transitional probabilities in an artificial language would scale up to the challenge of acquiring natural language. On the one hand, there is evidence that infants track transitional probabilities between syllables in highly controlled natural language input (e.g., Jusczyk et al. 1999, Pelucchi et al. 2009), and possibly even in their everyday language exposure (Ngon et al. 2013). On the other hand, computational studies (Yang 2004) and carefully controlled experiments using slightly more naturalistic artificial languages (e.g., Johnson & Tyler 2010) question the feasibility of statistical learning for word segmentation (see Johnson 2012 for a discussion). For example Johnson & Tyler (2010) presented 5- and 8-month-old Dutch-learning infants with one of two types of artificial languages. In one condition, infants heard a language containing four words of uniform length. In the other condition, infants heard a language containing four words with different lengths. The transitional probabilities between words were held constant across the two languages. Both age groups succeeded in segmenting words from the uniform-length language, but neither group succeeded with the mixed-length language. Thus, the authors concluded that infants' ability to track transitional probabilities between syllables might not scale up to the challenge of natural language, where word lengths are never perfectly uniform (see also Mersad & Nazzi 2012). And other work has suggested that, given natural language input, infants may rely more on acoustic-phonetic cues to word boundaries than on transitional probabilities (e.g., Johnson 2003). Questions regarding the ecological validity of statistical learning explanations for word segmentation are likely to continue in the years to come. A possibility consistent with all of the current data on both sides of this debate is that infants indeed track transitional probabilities between syllables in natural language, but not to the extent that they can rely solely on this information to bootstrap language-specific segmentation strategies.

4.1.3. Universal prosodic cues to word boundaries. A third approach that infants might use to find words involves universal prosodic cues to word boundaries. Recall our example of listening to a foreign language and having the impression that all of the words run together; contrary to this compelling impression, there are in fact some fully reliable cues to word boundaries (e.g., Endress & Hauser 2010). Speakers of every language pause between utterances (how else would they breathe?), and these pauses provide reliable cues to word boundaries. The proposal that infants might use utterance boundaries to learn about word boundaries has been termed the Edge Hypothesis (Seidl & Johnson 2006), and substantial evidence in support of this notion exists. First, corpus studies have demonstrated that speech directed to infants has a disproportionately high number of words flanked by utterance boundaries (e.g., Johnson et al. 2013, van de Weijer 1998), that frequent nouns tend to occur at utterance boundaries (Johnson et al. 2014), and that mothers highlight words of interest by aligning them with utterance boundaries (Aslin et al. 1996). Second, computational models have shown that increasing utterance boundary frequency improves segmentation performance (Frank et al. 2010) and that listeners may be able to use utterance edges to help bootstrap language-specific word segmentation strategies (Brent & Cartwright 1996; see Daland & Pierrehumbert 2011 for related work on the use of phrase boundaries). Third, adult artificial language studies have shown that listeners learn phonotactic patterns better when they occur at utterance edges (Endress & Mehler 2010; see also Slobin 1973 for the acquisition principle "pay attention to the ends of things") and suggest that utterance boundaries provide a much more efficient segmentation strategy than do transitional probabilities between syllables (Sohail & Johnson, forthcoming). Finally, experiments have shown that infants segment words from speech more readily when they are aligned with utterance boundaries than when they occur utterance medially (Johnson et al. 2014; Seidl & Johnson 2006, 2008).

The Edge Hypothesis is just one example of a universal prosodic cue to word boundaries. Other prosodic cues to word boundaries also play an important role in infants' early segmentation attempts, including the use of major phrase or clause boundaries to constrain lexical searches (e.g., Shukla et al. 2007, 2011), constraints on minimal word lengths such that all words contain at least one vowel (e.g., Brent & Cartwright 1996, Johnson et al. 2003), and possibly the Unique Stress Constraint (no word can contain more than one syllable with primary stress; Yang 2004). Of these proposed constraints, behavioral data so far certainly support infants' use of major phrase or clause boundaries and the implementation of a minimal word-length constraint. These strategies also fit well with what we know about newborns' perception of the speech signal (e.g., newborns are highly sensitive to prosody).

4.1.4. Summary of different word-finding strategies. In Section 4.1, I discuss how children first begin finding word forms to add to their proto-lexicon. We know that even very young infants possess language-specific strategies for finding words in speech, but how did they learn these strategies? It seems that infants need a language-general strategy for extracting at least a small cohort of words from speech before they can work out the language-specific segmentation strategies used by adult speakers of the language. Above, I outline several possible strategies that infants might use to find this initial cohort of words in speech (isolated words, distributional learning, and the use of universal prosody). All of these strategies probably play at least some role in infants' early segmentation strategies; however, a consensus on how this happens has yet to be achieved (e.g., Endress & Mehler 2009, Hay et al. 2011, Johnson & Tyler 2010, Yang 2004). In the future, an important factor in adjudicating between competing explanations for how infants first develop language-specific word segmentation skills might be additional research with infants learning languages that are structured very differently from English, such as Mandarin or Hungarian (for related discussion, see, e.g., Gervain & Mehler 2010, Johnson 2012, Nazzi et al. 2014, Peters 1981, Yang 2004).

4.2. Making Word Forms Meaningful

Extracting word forms from speech is an essential prerequisite to forming word–object pairings, but how do infants pair these word forms with the appropriate meaning? That is, once children have determined that 'Pinocchio' is a single word form (instead of, for example, four monosyllabic words), how do they work out what this word refers to? How do they know that 'Pinocchio' refers to the wooden boy rather than his cat or the whale that swallowed the boy? We know that older children have many word-learning cues at their disposal, including grammatical cues (e.g., Naigles 1990; Nappa et al. 2009; Paquette-Smith & Johnson, forthcoming), and various word-learning heuristics, including the mutual exclusivity principle (Markman 1990), but these strategies are not available to 6- to 9-month-old infants (see Golinkoff et al. 2000 for a review).

A proposed explanation for how infants learn words is that they track the statistical relationship between all of the word forms they hear and the objects they see in the world (e.g., Smith et al. 2014). By noting which word forms co-occur with which objects, infants may deduce form–referent mappings. As single labeling events can be ambiguous (because there are multiple possible referents to attach a word form to), proponents of this view have suggested that infants track these relationships across multiple situations. For example, imagine a child hears 'ball' while viewing a ball, a brush, a cup, and a table. At this point, the child has no way to determine which object the label refers to. But then later, the child might hear 'ball' again while viewing a ball, a spoon, and a sibling. This time, by comparing the objects present on the two occasions 'ball' was uttered, a child could deduce that 'ball' refers to the round bouncy thing. Artificial language studies have provided support for this type of cross-situational learning of words (e.g., Smith & Yu 2008).

However, much as in the debate over whether or not word forms can be extracted from speech by tracking transitional probabilities between syllables, there has been some disagreement over whether learning word–object associations through cross-situational statistics can scale up to the complexities of real-world language input (Medina et al. 2011, Smith et al. 2014, Yurovsky et al. 2013). For example, some researchers have argued that children do not simultaneously track all of the word forms they hear and all of the objects they see; rather, they form hypotheses about what words mean and then revise their hypotheses only when very clear evidence to the contrary is available (Trueswell et al. 2013). Researchers have also proposed a number of constraints, such as multimodal cues in parent–child interactions (e.g., Gogate 2010, Gogate & Hollich 2010, Jesse & Johnson 2012, Yu & Ballard 2007) or integration with grammatical knowledge (e.g., Hochmann et al. 2010, Monaghan & Mattock 2012), that may also help limit the number of word–object pairings children consider. It seems that a major focus of future research in this area will be on better understanding what sorts of cross-situational information are available to infants in the real-world complex visual and auditory scenes, and on how children integrate this information with other cues to word meaning.

4.3. Factors Affecting Word Learning

Various factors affect the acquisition of word form and word–object pairings in infancy (Werker & Curtin 2005). For example, young infants appear to form more robust acoustic-phonetic representations of words that occur highly frequently in the input, facilitating recognition of these items across acoustically distinct pronunciations (Singh et al. 2008). A similar pattern is seen in word learning, where toddlers find it easier to form word–object pairings when the label for the object is a familiar rather than unfamiliar word form (Swingley 2007). Word forms that are consonant initial are segmented from speech before those that are vowel initial (Kim & Sundara 2014, Seidl & Johnson 2008), and word–object pairings are formed more readily when labels are composed of legal phonotactic sequences or frequent lexical stress patterns (Graf Estes & Bowen 2013). Prosodic characteristics (e.g., Seidl & Johnson 2006, Shukla et al. 2011) and grammatical word class (e.g., Gillette et al. 1999, Nazzi et al. 2005, Willits et al. 2014) also affect the ease of acquisition, as does speech register (Ma et al. 2011, Thiessen et al. 2005). Perhaps relatedly, hearing many variable tokens of a word can also aid the formation of word–object pairings (Rost & McMurray 2009).

Visual factors affect word learning too. For example, young infants readily form word–object pairings in the lab only if each labeling of an object is accompanied by a synchronous gesture (Gogate 2010), or if the labeled object is particularly interesting to look at (Pruden et al. 2006). Finally, in toddlerhood, word–object pairings are learned better when taught via socially contingent interactions (Roseberry et al. 2014).

4.4. How Word Forms, Word–Object Pairings, and "Real" Words Relate

In Section 4, I classify both the extraction of word forms from speech and the formation of word–object pairings as different aspects of the overall process of learning a word. These issues have often been discussed separately in the literature, with researchers focusing either on how infants find word forms in speech or how infants form word–object pairings. Indeed, extracting words from speech is often treated in the literature as different from word learning. However, more recently, language researchers have become increasingly aware that word form learning and word–object pairings are tightly linked, and that these two learning processes should not be studied independently of one another (e.g., Graf Estes et al. 2013).

My discussion of word meaning has also focused on the formation of word–object pairings, largely avoiding the issue of how word–object pairings relate to "real" words. Researchers do not always agree on whether to classify early "associations" between word forms and the objects they refer to as "real" words or as simple nonsymbolic proto-words (Bloom 2001, Nazzi & Bertoncini 2003, Sloutsky & Fisher 2004, Waxman & Gelman 2009). Some research suggests that infants' early word–object pairings are quite real (e.g., Fulkerson & Waxman 2007), but this issue is still controversial. New methodological tools have been developed that could be used to help map out how fully young children understand words (i.e., whether they are genuine words embedded within a categorical semantic framework rather than mere associations; Arias-Trejo & Plunkett 2013, Johnson et al. 2011, Wojcik & Saffran 2015), but at this point many questions regarding the representational nature of early words remain to be solved.

5. LINKING INDIVIDUAL VARIATION IN INFANCY TO LANGUAGE OUTCOMES

Infants must learn the phonological structure of their native language to learn words, and learning words in turn helps infants further fine-tune their understanding of other aspects of language structure. Importantly, this process does not seem to depend on being spoon-fed predigested bite-sized bits of language by one's parents (e.g., Aslin et al. 1996, van de Weijer 1998), but it does appear to depend on positive social interactions (e.g., Bloom et al. 1987, Goldstein & Schwade 2008) and the quality and quantity of language input received by the child (e.g., Cartmill et al. 2013, Hart & Risley 1995, Weisleder & Fernald 2013).

Recently, a growing literature has examined the relationship between early experiences, the development of language-specific phonological knowledge, and long-term language outcomes (see Cristia et al. 2014 for a review). Live social interaction (as opposed to off-line videotaped interactions) appears to facilitate infants' acquisition of sound structure knowledge (Kuhl et al. 2003), and there is a positive relationship between the clarity of a mother's speech and her child's speech perception skills (Liu et al. 2003). Moreover, the earlier children learn to ignore phonetic contrasts that do not signal meaningful differences in the native language, the better their subsequent language development (Kuhl et al. 2008). There is also evidence that the early development of word segmentation abilities in infancy is linked to greater vocabulary skills several years later (Junge & Cutler 2014, Newman et al. 2006, Singh et al. 2012). Taken together, the results from these studies underscore the importance of early phonological development and word form learning for subsequent language development.

One area for future research might be to test whether there is a relationship between attunement to the native language phonetic inventory and the development of word segmentation skills in infancy, and to ask whether these two skills make independent contributions to subsequent language development. It would also be interesting to broaden the range of later language skills that are examined in relation to early speech perception development. For example, comprehension of accented speech has been argued to require phonological constancy (i.e., recognition of a speech segment across natural variation in its phonetic realization), which has been argued to emerge only at approximately 19 months of age (e.g., Mulak et al. 2013; see, however, van Heugten & Johnson 2014). Does either attunement to native language phonetic categories or the development of word segmentation abilities in infancy predict how quickly children develop the ability to cope with accented speech? By addressing questions like these, researchers could begin to sharpen our understanding of how language input, the acquisition of language-specific sound structure, word segmentation abilities, and the development of subsequent language skills are linked. This, in turn, could help researchers unify models of word learning and phonological development into a single more comprehensive model of early language acquisition.

In addition to uncovering links between infants' performance in speech perception tasks and their subsequent language development, researchers are beginning to discover neural predictors of language development. For example, Dutch infants produce brain responses to familiarized words heard in speech nearly 2 months earlier than they produce any outward behavioral evidence of segmenting words from speech, and these neural responses are predictive of language development at the age of 3 years (Kooijman et al. 2013). Future studies combining physiological and behavioral measures of word segmentation hold great promise for further understanding the complex relationship between word form learning and other aspects of language acquisition (Kooijman et al. 2008).

Other methodological advances needed to advance our understanding of early language development include the creation of testing procedures that are sensitive enough to detect individual variation in children's perceptual development. At the moment, nearly all research examining early speech perception capabilities involves collapsing across data collected from many children. It is possible that a more nuanced approach to early perceptual development could reveal that children employ different strategies for extracting linguistically relevant information from the speech signal—that is, by averaging across the performance of many children, current infant speech testing methodologies may be masking individual variation in early language-learning strategies (e.g., although most children may tend to follow the stereotypical pattern of development described in the literature, some children may have alternative strategies such as focusing their attention on whole word forms or phrasal contours; for a related discussion in the production literature, see Peters 1977).

6. FITTING INFANT WORD LEARNING INTO THE BIGGER PICTURE

By now, it should be clear how infants' early understanding of the native language sound structure facilitates word form learning (and thus, eventually, word learning). But how does early word learning contribute to other aspects of language development? Is there a reciprocal feedback loop between word form learning, phonological development, and other areas of language development? In this section, I first discuss the relationship between word form learning and infants' acquisition of the phonetic categories of their native language. I then briefly discuss how development in these two areas relates to the acquisition of grammatical structure.

6.1. Words and Phones

Section 2 describes how children become tuned to the native language in their first year of life. By roughly 10 to 12 months, infants probably already recognize the sound patterns of hundreds of word forms (Swingley 2009), and begin showing increased sensitivity to phoneme contrasts that occur in the language and decreased sensitivity to phoneme contrasts that do not occur in the language (Kuhl et al. 2008; but see Tyler et al. 2014). At the same time, infants also begin showing dramatic improvements in their ability to segment word forms from speech (e.g., Jusczyk et al. 1999, Kim & Sundara 2014) and are better able to recognize word forms across acoustic-phonetic variation (e.g., Houston & Jusczyk 2000). Are the simultaneous improvements in infants' phonetic category knowledge and word-recognition abilities a coincidence, or are improvements in these two areas somehow linked? In other words, what is the relationship between word learning and the acquisition of language-specific phonetic categories?

One could imagine that children (much like field linguists) acquire the phoneme inventory of their language only after learning a sizeable number of minimal pairs (if 'pat' and 'bat' are different words, than /p/ and /b/ must be phonemes in the language). However, infants display attunement to the native language phonology by age 10 to 12 months, well before they could possibly understand enough minimal pairs to sustain this sort of learning strategy. A more recent proposal to explain how infants acquire the phonetic category learning involves tracking the statistical distribution of sounds in the native language (Maye et al. 2002). The advantage of this approach over the minimal pair learning approach is that it depends entirely on bottom-up acoustic-phonetic information, and requires no word (or word form) knowledge. Evidence for the feasibility of this approach has been provided by infant behavioral studies. Maye et al. created a voice-onset time (VOT) continuum between [da] and [ta], and infants were familiarized with either a bimodal (lots of [da]- and [ta]-like sounds, but few sounds in between) or a unimodal distribution of the continuum (lots of sounds that were between [da] and [ta]). In a discrimination test following familiarization, infants in the bimodal group distinguished [da] from [ta], whereas infants in the unimodal group did not. The remarkable success of this study led to the adoption of this mechanism in several prominent models of infant speech development (e.g., Werker & Curtin 2005, Kuhl et al. 2008). However, much like statistical learning for word segmentation and cross-situational statistics for learning word–object pairings, some researchers have questioned whether this approach would work very efficiently with real-world natural language input (e.g., Yeung & Werker 2009).

Although learning minimal pairs cannot explain how children acquire the phoneme inventory in their language, this does not necessarily mean that word (or word form) knowledge plays no role at all in this process. Perhaps infants' acquisition of words and their acquisition of phoneme categories proceed hand in hand, with success in each domain feeding into the other. Language researchers differ greatly on the details of how this process could work (e.g., Feldman et al. 2013, Martin et al. 2013, Yeung & Werker 2009), but a vaguely specified generic version might be as follows. Early on, infants start pulling out word-sized chunks from speech. Some of these word-sized chunks are associated with meaning, and some are not. Regardless, all of these word-sized chunks (meaningless or not) help infants work out the typical sound structure of words in the native language, which in turn allows infants to pull out further words as well as additional tokens of already-known word forms. As the size and robustness of infants' stock of word forms grow, infants are also tracking information about the distribution of sounds in these words. Importantly, according to this view, infants are not merely tracking the frequency of different sounds (as proposed by Maye et al. 2002); they are also tracking the distributions of sounds in relation to word forms. By linking sound distributions to words forms in the proto-lexicon, infants may have additional information to help them bootstrap the phonetic categories of the native language from speech. As infants learn more about the sound structure of their language, it becomes easier to extract phonological details from speech while trying to simultaneously work out word meanings. Thus, it becomes easier to map word forms to meaning. In short, according to this view, word learning and phonological development can still be integrally related without the need for minimal pairs in children's early vocabularies.

6.2. Sound Structure, Word Forms, and Grammatical Structure

The word learning literature has historically been more focused on children's acquisition of content words (e.g., nouns and verbs) than on closed-class function words (e.g., pronouns and determiners). One reason may be that children often omit function words from their early speech (e.g., Brown 1973), so researchers thought children might acquire these items late (see Gerken et al. 1990 for a discussion). However, if children's failure to produce function words were an indication of their failure to perceive or recognize them, then this would have serious implications for children's acquisition of syntactic structure. More recent research has suggested that children start learning the correct positioning (see, e.g., Gerken 2002 for a review) and meaning (Saylor et al. 2011; see also Hochmann et al. 2010 for a related discussion) of function words early, and that attention to function words helps infants expand their vocabularies (e.g., Shi et al. 2006) and discover the grammatical class of words forms (Chemla et al. 2009, Höhle et al. 2004, Shi & Melançon 2010). Attention to the distribution of function words has even been argued to help 8-month-olds learn the ordering of words in the native language (Gervain & Mehler 2010, Gervain et al. 2008). Thus, infants' precocious understanding of function words suggest that word learning and grammatical development proceed hand in hand from early on in development, just as do word learning and phonological development.

7. CONCLUSIONS AND FUTURE DIRECTIONS

Language acquisition begins in the womb. In the latter half of the first year of life, infants acquire many word–object pairings and begin using language-specific knowledge to extract new word forms from speech. By age 10 to 12 months, infants have typically produced their first words and language experience has shaped the way infants attend to phonetic contrasts. At this point, infants already recognize a large number of word forms, and are beginning to use closed-class function words to work out the structure of their language. Between the ages of 7 and 18 months, infants' ability to deal with acoustic-phonetic variation in the realization of words improves, as does their rate of learning new words. And throughout this entire process, infants understand far more than they say.

It seems that children are simultaneously learning the sound structure of their native language, building a proto-lexicon, and beginning to work out the grammatical structure of their language. Learning at each of these levels depends on learning occurring at the other levels, such that the key to unlocking the linguistic structure of language lies in the integration of information across these domains. But how exactly is information integrated across domains? Can the word-learning strategies outlined in this review work equally well for all of the world's languages? How strongly constrained are infants' language-learning mechanisms? What is the precise relationship between phoneme acquisition and lexical development? When do infants' lexical representations become abstract? What is the relationship between early word–object pairings and "real" words? How do different learning environments (e.g., multilingualism, atypical social interactions) affect phonological development? Do perception studies that average across many participants mask important differences in children's language-learning styles? How can the relationship between perception and production in early language development best be understood? Current models and data provide only partial answers to these questions. As the field continues to grow and advance, we look forward to the development of exciting and innovative new models that can successfully integrate all of these factors into a unified model of language acquisition.

SUMMARY POINTS

1. Language acquisition begins in the womb, where infants receive exposure to the rhythm and melody of their mothers' native tongue.
2. By 6 months of age, infants' proto-lexicons already contain many word forms and word–object pairings. By the time children produce their first word at approximately 1 year of age, they may already recognize the sound patterns of hundreds of word forms.
3. While building a proto-lexicon, children are also acquiring the phoneme inventory of the native language and beginning to learn about how grammatical sentences are formed. Integration of information across these domains appears to be the key to infants' success at unlocking the linguistic structure of the native language.
4. Future research will need to further investigate how infants acquire their first words, what role lexical development plays in the sharpening of phonetic boundaries, how much infants can learn about the structure of their native language from statistics alone, how infants cope with variation in the speech signal, how neural development relates to language development, and whether different learning styles can be observed in early speech perception.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank Helen Buckler, Anne Cutler, and Marieke van Heugten for constructive feedback on a previous version of this manuscript. Manuscript preparation was supported by funding from the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

Altvater-Mackensen N, Mani N. 2013. Word-form familiarity bootstraps infant speech segmentation. *Dev. Sci.* 16:980–90

Anderson JL, Morgan JL, White KS. 2003. A statistical basis for speech sound discrimination. *Lang. Speech* 46:155–82

Arias-Trejo N, Plunkett K. 2013. What's in a link? Associative and taxonomic priming effects in the infant lexicon. *Cognition* 128:214–27

Aslin RN, Shukla M, Emberson LL. 2015. Hemodynamic correlates of cognition in human infants. *Annu. Rev. Psychol.* 66:349–79

Aslin RN, Werker JF, Morgan JL. 2002. Innate phonetic boundaries revisited. *J. Acoust. Soc. Am.* 112:1257–60

Aslin RN, Woodward J, LaMendola N, Bever T. 1996. Models of word segmentation in fluent maternal speech to infants. In *Signal to Syntax*, ed. JL Morgan, K Demuth, pp. 117–34. New York: Psychology

Batchelder EO. 2002. Bootstrapping the lexicon: a computational model of infant speech segmentation. *Cognition* 83:167–206
Bergelson E, Swingley D. 2012. At 6–9 months, human infants know the meanings of many common nouns. *PNAS* 109:3253–58
Bloom K, Russell A, Wassenberg K. 1987. Turn taking affects the quality of infant vocalizations. *J. Child Lang.* 14:211–27
Bloom P. 2001. Précis of "How children learn the meanings of words." *Behav. Brain Sci.* 24:1095–103
Bortfeld H, Morgan JL, Golinkoff RM, Rathbun K. 2005. Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychol. Sci.* 16:298–304
Bouchon C, Floccia C, Fux T, Adda-Decker M, Nazzi T. 2014. Call me Alix, not Elix: Vowels are more important than consonants in own name recognition at 5 months. *Dev. Sci.* 18:587–98
Bowerman M. 1976. Semantic factors in the acquisition of rules for word use and sentence construction. In *Directions in Normal and Deficient Language Development*, ed. D Morehead, A Morehead, pp. 99–179. Baltimore: Univ. Park Press
Brent MR, Cartwright T. 1996. Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61:93–125
Brent MR, Siskind JM. 2001. The role of exposure to isolated words in early vocabulary development. *Cognition* 81:33–44
Brown R. 1973. *A First Language: The Early Stages*. Cambridge, MA: Harvard Univ. Press
Burnham DK. 1986. Developmental loss of speech perception: exposure to and experience with a first language. *Appl. Psycholinguist.* 7:207–39
Cairns GF, Butterfield EC. 1975. Assessing infant's auditory functioning. In *Exceptional Infant*, vol. 3: *Assessment and Intervention*, ed. BF Friedlander, pp. 84–108. New York: Brunner/Mazel
Cartmill EA, Armstrong BF, Gleitman LR, Goldin-Meadow S, Medina TN, Trueswell JC. 2013. Quality of early parent input predicts child vocabulary 3 years later. *PNAS* 110:11278–83
Chemla E, Mintz TH, Bernal S, Christophe A. 2009. Categorizing words using "frequent frames": what cross-linguistic analyses reveal about distributional acquisition strategies. *Dev. Sci.* 12:396–406
Choi J. 2014. *Rediscovering a Forgotten Language*. Nijmegen, Neth.: Max Planck Inst. Psycholinguist.
Christophe A, Guasti T, Nespor M, Dupoux E, Van Ooyen B. 1997. Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Lang. Cogn. Process.* 12:585–612
Cole RA, Jakimik JA. 1980. A model of speech perception. In *Perception and Production of Fluent Speech*, ed. RA Cole, pp. 133–42. Hillsdale, NJ: Erlbaum
Cooper RP, Aslin RN. 1990. Preference for infant-directed speech in the first month after birth. *Child Dev.* 61:1584–95
Cristia A, Seidl A, Junge C, Soderstrom M, Hagoort P. 2014. Predicting individual variation in language from infant speech perception measures. *Child Dev.* 85:1330–45
Cutler A. 2012. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge, MA: MIT Press
Cutler A, Mehler J. 1993. The periodicity bias. *J. Phon.* 21:103–8
Cutler A, Norris D. 1988. The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Hum. Percept. Perform.* 14:113–21
Dahan D, Brent MR. 1999. On the discovery of novel wordlike units from utterances: an artificial-language study with implications for native-language acquisition. *J. Exp. Psychol. Gen.* 128:165–85
Daland R, Pierrehumbert JB. 2011. Learning diphone-based segmentation. *Cogn. Sci.* 35:119–55
DeCasper AJ, Spence MJ. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.* 9:133–50
Eimas PD, Siqueland ER, Jusczyk PW, Vigorito J. 1971. Speech perception in infants. In *First Language Acquisition: The Essential Readings*, ed. BC Lust, C Foley, pp. 279–84. Oxford, UK: Wiley Blackwell
Endress AD, Hauser MD. 2010. Word segmentation with universal prosodic cues. *Cogn. Psychol.* 61:177–99
Endress AD, Mehler J. 2009. The surprising power of statistical learning: when fragment knowledge leads to false memories of unheard words. *J. Mem. Lang.* 60:351–67
Endress AD, Mehler J. 2010. Perceptual constraints in phonotactic learning. *J. Exp. Psychol. Hum. Percept. Perform.* 36:235–50

Feldman NH, Griffiths TL, Goldwater S, Morgan JL. 2013. A role for the developing lexicon in phonetic category acquisition. *Psychol. Rev.* 120:751–78

Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, et al. 1994. Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* 59:174–85

Fernald A, Perfors A, Marchman VA. 2006. Picking up speed in understanding: speech processing efficiency and vocabulary growth across the 2nd year. *Dev. Psychol.* 42:98–116

Fernald A, Zangl R, Portillo AL, Marchman VA. 2008. Looking while listening: using eye movements to monitor spoken language comprehension by infants and young children. In *Developmental Psycholinguistics: On-Line Methods in Children's Language Processing*, ed. IA Sekerina, EM Fernández, H Clahsen, pp. 97–135. Amsterdam: Benjamins

Frank MC, Goldwater S, Griffiths T, Tenenbaum JB. 2010. Modeling human performance in statistical word segmentation. *Cognition* 117:107–25

Friedrich M, Friederici AD. 2011. Word learning in 6-month-olds: fast encoding—weak retention. *J. Cogn. Neurosci.* 23:3228–40

Fulkerson AL, Waxman SR. 2007. Words (but not tones) facilitate object categorization: evidence from 6- and 12-month-olds. *Cognition* 105:218–28

Gerken L. 2002. Early sensitivity to linguistic form. *Annu. Rev. Lang. Acquis.* 2:1–36

Gerken L, Landau B, Remez R. 1990. Function morphemes in young children's speech perception and production. *Dev. Psychol.* 26:204–16

Gervain J, Mehler J. 2010. Speech perception and language acquisition in the first year of life. *Annu. Rev. Psychol.* 61:191–218

Gervain J, Nespor M, Mazuka R, Horie R, Mehler J. 2008. Bootstrapping word order in prelexical infants: a Japanese–Italian cross-linguistic study. *Cogn. Psychol.* 57:56–74

Gillette J, Gleitman H, Gleitman L, Lederer A. 1999. Human simulations of vocabulary learning. *Cognition* 73:135–76

Gogate LJ. 2010. Learning of syllable–object relations by preverbal infants: the role of temporal synchrony and syllable distinctiveness. *J. Exp. Child Psychol.* 105:178–97

Gogate LJ, Hollich G. 2010. Invariance detection within an interactive system: a perceptual gateway to language development. *Psychol. Rev.* 117:496–516

Goldstein MH, Schwade JA. 2008. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* 19:515–23

Golinkoff RM, Hirsh-Pasek K, Bloom L, Smith LB, Woodward AL, et al. 2000. *Becoming a Word Learner: A Debate on Lexical Acquisition*. New York: Oxford Univ. Press

Graf Estes K, Bowen S. 2013. Learning about sounds contributes to learning about words: effects of prosody and phonotactics on infant word learning. *J. Exp. Child Psychol.* 114:405–17

Graf Estes K, Evans JL, Alibali MW, Saffran JR, Estes KG. 2013. Can infants map meaning to words in newly segmented words? *Psychol Sci.* 18:254–60

Harris ZS. 1955. From phoneme to morpheme. *Language* 31:190–222

Hart B, Risley TR. 1995. *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Brookes

Hay JF, Pelucchi B, Estes KG, Saffran JR. 2011. Linking sounds to meanings: infant statistical learning in a natural language. *Cogn. Psychol.* 63:93–106

Hochmann JR, Endress AD, Mehler J. 2010. Word frequency as a cue for identifying function words in infancy. *Cognition* 115:444–57

Höhle B, Weissenborn J, Kiefer D, Schulz A, Schmitz M. 2004. Functional elements in infants' speech processing: the role of determiners in the syntactic categorization of lexical elements. *Infancy* 5:341–53

Houston DM, Jusczyk PW. 2000. The role of talker-specific information in word segmentation by infants. *J. Exp. Psychol. Hum. Percept. Perform.* 26:1570–82

Houston DM, Jusczyk PW, Kuijpers C, Coolen R, Cutler A. 2000. Cross-language word segmentation by 9-month-olds. *Psychon. Bull. Rev.* 7:504–9

Jesse A, Johnson EK. 2012. Prosodic temporal alignment of co-speech gestures to speech facilitate referent resolution. *J. Exp. Psychol. Hum. Percept. Perform.* 38:1567–81

Johnson EK. 2003. *Word segmentation during infancy: the role of subphonemic cues to word boundaries*. PhD thesis, Dep. Linguist., Johns Hopkins Univ., Baltimore

Johnson EK. 2012. Bootstrapping langauge: Are infant statisticians up to the job? In *Statistical Learning and Language Acquisition*, ed. P Rebuschat, J Williams, pp. 55–90. Boston: de Gruyter

Johnson EK, Jusczyk PW. 2001. Word segmentation by 8-month-olds: when speech cues count more than statistics. *J. Mem. Lang.* 44:548–67

Johnson EK, Jusczyk PW, Cutler A, Norris D. 2003. Lexical viability constraints on speech segmentation by infants. *Cogn. Psychol.* 46:65–97

Johnson EK, Lahey M, Ernestus M, Cutler A. 2013. A multimodal corpus of speech to infant and adult listeners. *J. Acoust. Soc. Am.* 134:EL534

Johnson EK, McQueen JM, Huettig F. 2011. Toddlers' language-mediated visual search: They need not have the words for it. *Q. J. Exp. Psychol.* 64:1672–82

Johnson EK, Seidl A, Tyler MD. 2014. The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PLOS ONE* 9:e83546

Johnson EK, Tyler MD. 2010. Testing the limits of statistical learning for word segmentation. *Dev. Sci.* 13:339–45

Johnson EK, Zamuner T. 2010. Using infant and toddler testing methods in language acquisition research. In *Experimental Methods in Language Acquisition Research*, ed. E Blom, S Unsworth, pp. 73–94. Amsterdam: Benjamins

Junge C, Cutler A. 2014. Early word recognition and later language skills. *Brain Sci.* 4:532–59

Jusczyk PW. 1997. *The Discovery of Spoken Language*. Cambridge, MA: MIT Press

Jusczyk PW, Aslin RN. 1995. Infants' detection of the sound patterns of words in fluent speech. *Cogn. Psychol.* 29:1–23

Jusczyk PW, Hohne EA. 1997. Infants' memory for spoken words. *Science* 277:1984–86

Jusczyk PW, Houston DM, Newsome M. 1999. The beginnings of word segmentation in English-learning infants. *Cogn. Psychol.* 39:159–207

Keren-Portnoy T, Majorano M, Vihman MM. 2009. From phonetics to phonology: the emergence of first words in Italian. *J. Child Lang.* 36:235–67

Kim YJ, Sundara M. 2014. Segmentation of vowel-initial words is facilitated by function words. *J. Child Lang.* 27:1–25

Kooijman V, Johnson EK, Cutler A. 2008. Reflections on reflections of infant word recognition. In *Early Language Development: Bridging Brain and Behaviour*, ed. AD Friederici, G Thierry, pp. 91–114. Amsterdam: Benjamins

Kooijman V, Junge C, Johnson EK, Hagoort P, Cutler A. 2013. Predictive brain signals of linguistic development. *Front. Psychol.* 4:25

Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M, Nelson T. 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. B* 363:979–1000

Kuhl PK, Meltzoff AN. 1996. Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* 100:2425–38

Kuhl PK, Tsao F-M, Liu H-M. 2003. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *PNAS* 100:9096–101

Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255:606–8

Labov W, Labov T. 1978. The phonetics of *cat* and *mama*. *Language* 54:816–52

Lecanuet J, Schaal B. 2002. Sensory performances in the human foetus: a brief summary of research. *Intellectica* 1:29–56

Lew-Williams C, Pelucchi B, Saffran JR. 2011. Isolated words enhance statistical language learning in infancy. *Dev. Sci.* 14:1323–29

Lidz J, Gagliardi A. 2015. How nature meets nurture: Universal Grammar and statistical learning. *Annu. Rev. Linguist.* 1:333–53

Liu HM, Kuhl PK, Tsao FM. 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Dev. Sci.* 6:1–10

Ma W, Golinkoff RM, Houston D, Hirsh-Pasek K. 2011. Word-learining in infant- and adult-directed speech. *Lang. Learn. Dev.* 7:209–25

Mampe B, Friederici AD, Christophe A, Wermke K. 2009. Newborns' cry melody is shaped by their native language. *Curr. Biol.* 19:1994–97

Mandel DR, Jusczyk PW, Pisoni DB. 1995. Infants' recognition of the sound patterns of their own names. *Psychol. Sci.* 6:314–17

Markman E. 1990. Constraints children place on word meanings. *Cogn. Sci.* 14:55–77

Martin A, Peperkamp S, Dupoux E. 2013. Learning phonemes with a proto-lexicon. *Cogn. Sci.* 37:103–24

Mattock K, Molnar M, Polka L, Burnham D. 2008. The developmental course of lexical tone perception in the first year of life. *Cognition* 106:1367–81

Mattys SL, Jusczyk PW. 2001. Phonotactic cues for segmentation of fluent speech by infants. *Cognition* 78:91–121

Maye J, Werker JF, Gerken L. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82:B101–11

McMurray B. 2007. Defusing the childhood vocabulary explosion. *Science* 317:631

McQueen JM. 1998. Segmentation of continuous speech using phonotactics. *J. Mem. Lang.* 39:21–46

Medina TN, Snedeker J, Trueswell JC, Gleitman LR. 2011. How words can and cannot be learned by observation. *PNAS* 108:9014–19

Mersad K, Nazzi T. 2012. When mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. *Lang. Learn. Dev.* 8:303–15

Monaghan P, Mattock K. 2012. Integrating constraints for learning word–referent mappings. *Cognition* 123:133–43

Moon C, Cooper RP, Fifer WP. 1993. Two-day-olds prefer their native language. *Infant Behav. Dev.* 16:495–500

Moon CM, Lagercrantz H, Kuhl PK. 2013. Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatr.* 102:156–60

Mulak KE, Best CT, Tyler MD, Kitamura C, Irwin JR. 2013. Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Dev.* 84:2064–78

Naigles L. 1990. Children use syntax to learn verb meanings. *J. Child Lang.* 17:357–74

Nappa R, Wessel A, McEldoon KL, Gleitman LR, Trueswell JC. 2009. Use of speaker's gaze and syntax in verb learning. *Lang. Learn. Dev.* 5:203–34

Narayan CR, Werker JF, Beddor PS. 2010. The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Dev. Sci.* 13:407–20

Nazzi T, Bertoncini J. 2003. Before and after the vocabulary spurt: two modes of word acquisition? *Dev. Sci.* 6:136–42

Nazzi T, Bertoncini J, Mehler J. 1998. Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24:756–66

Nazzi T, Dilley LC, Jusczyk AM, Shattuck-Hufnagel S, Jusczyk PW. 2005. English-learning infants' segmentation of verbs from fluent speech. *Lang. Speech* 48:279–98

Nazzi T, Mersad K, Sundara M, Iakimova G, Polka L. 2014. Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *J. Child Lang.* 41:600–633

Newman R, Ratner NB, Jusczyk AM, Jusczyk PW, Dow KA. 2006. Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Dev. Psychol.* 42:643–55

Ngon C, Martin A, Dupoux E, Cabrol D, Dutat M, Peperkamp S. 2013. (Non)words, (non)words, (non)words: evidence for a protolexicon during the first year of life. *Dev. Sci.* 16:24–34

Paquette-Smith M, Johnson EK. Toddlers' use of grammatical and social cues to learn novel words. *Lang. Learn. Dev.* Forthcoming

Partanen E, Kujala T, Tervaniemi M, Huotilainen M. 2013. Prenatal music exposure induces long-term neural effects. *PLOS ONE* 8:e78946

Patterson ML, Werker JF. 2003. Two-month-old infants match phonetic information in lips and voice. *Dev. Sci.* 6:191–96

Pelucchi B, Hay JF, Saffran JR. 2009. Statistical learning in a natural language by 8-month-old infants. *Child Dev.* 80:674–85
Peña M, Maki A, Kovacić D, Dehaene-Lambertz G, Koizumi H, et al. 2003. Sounds and silence: an optical topography study of language recognition at birth. *PNAS* 100:11702–5
Peters AM. 1977. Language learning strategies: Does the whole equal the sum of the parts? *Language* 53:560–73
Peters AM. 1981. Language typology and the segmentation problem in early child language acquisition. In *Proceedings of the 7th Annual Meeting of the Berkeley Linguistics Society*, pp. 236–48. Washington, DC: Linguist. Soc. Am.
Polka L, Werker JF. 1994. Developmental changes in perception of nonnative vowel contrasts. *J. Exp. Psychol. Hum. Percept. Perform.* 20:421–35
Pruden SM, Hirsh-Pasek K, Golinkoff RM, Hennon EA. 2006. The birth of words: Ten-month-olds learn words through perceptual salience. *Child Dev.* 77:266–80
Rescorla L. 1980. Overextension in early language development. *J. Child Lang.* 2:321–35
Romberg AR, Saffran JR. 2010. Statistical learning and language acquisition. *Wiley Interdiscip. Rev. Cogn. Sci.* 1:906–14
Roseberry S, Hirsh-Pasek K, Golinkoff RM. 2014. Skype me! Socially contingent interactions help toddlers learn language. *Child Dev.* 85:956–70
Rost GC, McMurray B. 2009. Speaker variability augments phonological processing in early word learning. *Dev. Sci.* 12:339–49
Rytting CA, Brew C, Fosler-Lussier E. 2010. Segmenting words from natural speech: subsegmental variation in segmental cues. *J. Child Lang.* 37:513–43
Saffran J, Werker JF, Werner L. 2006. The infant's auditory world: hearing, speech, and the beginnings of language. *Handb. Child Dev.* 6:58–108
Saffran JR, Aslin RN, Newport EL. 1996. Statistical learning by 8-month-olds. *Science* 274:1926–28
Sahni SD, Seidenberg MS, Saffran JR. 2010. Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Dev.* 81:727–36
Saylor MM, Ganea PA, Vázquez MD. 2011. What's mine is mine: Twelve-month-olds use possessive pronouns to identify referents. *Dev. Sci.* 14:859–64
Schmale R, Cristià A, Seidl A, Johnson EK. 2010. Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy* 15:650–62
Seidl A, Johnson EK. 2006. Infant word segmentation revisited: Edge alignment facilitates target extraction. *Dev. Sci.* 9:565–73
Seidl A, Johnson EK. 2008. Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *J. Child Lang.* 35:1–24
Shi R, Cutler A, Werker JF, Cruickshank M. 2006. Frequency and form as determinants of functor sensitivity in English-acquiring infants. *J. Acoust. Soc. Am.* 119:EL61–67
Shi R, Melançon A. 2010. Syntactic categorization in French-learning infants. *Infancy* 15:517–33
Shukla M, Nespor M, Mehler J. 2007. An interaction between prosody and statistics in the segmentation of fluent speech. *Cogn. Psychol.* 54:1–32
Shukla M, White KS, Aslin RN. 2011. Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *PNAS* 108:6038–43
Shultz S, Vouloumanos A, Bennett RH, Pelphrey K. 2014. Neural specialization for speech in the first months of life. *Dev. Sci.* 5:766–74
Singh L, Liederman J, Mierzejewski R, Barnes J. 2011. Rapid reacquisition of native phoneme contrasts after disuse: You do not always lose what you do not use. *Dev. Sci.* 14:949–59
Singh L, Morgan JL, White KS. 2004. Preference and processing: the role of speech affect in early spoken word recognition. *J. Mem. Lang.* 51:173–89
Singh L, Nestor SS, Bortfeld H. 2008. Overcoming the effects of variation in infant speech segmentation: influences of word familiarity. *Infancy* 13:57–74
Singh L, Reznick SJ, Xuehua L. 2012. Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Dev. Sci.* 15:482–95
Skoruppa K, Pons F, Christophe A, Bosch L, Dupoux E, et al. 2009. Language-specific stress perception by 9-month-old French and Spanish infants. *Dev. Sci.* 12:914–19

Slobin DI. 1973. Cognitive prerequisites for acquisition of grammar. In *Studies of Child Language Development*, ed. CA Ferguson, DJ Slobin, pp. 175–208. New York: Holt, Rinehart & Winston

Sloutsky VM, Fisher AV. 2004. Induction and categorization in young children: a similarity-based model. *J. Exp. Psychol. Gen.* 133:166–88

Smith L, Yu C. 2008. Infants rapidly learn word–referent mappings via cross-situational statistics. *Cognition* 106:1558–68

Smith LB, Suanda SH, Yu C. 2014. The unrealized promise of infant statistical word–referent learning. *Trends Cogn. Sci.* 18:251–58

Sohail J, Johnson EK. How transitional probabilities and the edge effect contribute to listeners' phonological bootstrapping success. *Lang. Learn. Dev.* Forthcoming

Styles S, Plunkett K. 2009. What is "word understanding" for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *J. Child Lang.* 36:895–908

Swingley D. 2005. Statistical clustering and the contents of the infant vocabulary. *Cogn. Psychol.* 50:86–132

Swingley D. 2007. Lexical exposure and word-form encoding in 1.5-year-olds. *Dev. Psychol.* 43:454–64

Swingley D. 2009. Contributions of infant word learning to language development. *Philos. Trans. R. Soc. B* 364:3617–32

Thiessen ED, Erickson LC. 2013. Discovering words in fluent speech: the contribution of two kinds of statistical information. *Front. Psychol.* 3:590

Thiessen ED, Hill EA, Saffran JR. 2005. Infant-directed speech facilitates word segmentation. *Infancy* 7:53–71

Thiessen ED, Saffran JR. 2003. When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev. Psychol.* 39:706–16

Thiessen ED, Saffran JR. 2007. Learning to learn: infants' acquisition of stress-based strategies for word segmentation. *Lang. Learn. Dev.* 3:73–100

Tincoff R, Jusczyk PW. 1999. Some beginnings of word comprehension in 6-month-olds. *Psychol. Sci.* 10:172–75

Tincoff R, Jusczyk PW. 2012. Six-month-olds comprehend words that refer to parts of the body. *Infancy* 17:432–44

Tomasello M, Mervis CB. 1994. The instrument is great, but measuring comprehension is still a problem. *Monogr. Soc. Res. Child Dev.* 59:174–79

Trehub SE. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Dev.* 47:466–72

Trueswell JC, Medina TN, Hafri A, Gleitman LR. 2013. Propose but verify: Fast mapping meets cross-situational word learning. *Cogn. Psychol.* 66:126–56

Tyler MD, Best CT, Goldstein LM, Antoniou M. 2014. Investigating the role of articulatory organs and perceptual assimilation of native and non-native fricative place contrasts. *Dev. Psychobiol.* 56:210–27

van de Weijer JC. 1998. *Language Input for Word Discovery*. Nijmegen, Neth.: Max Planck Inst. Psycholinguist.

van Heugten M, Johnson EK. 2012. Infants exposed to fluent natural speech succeed at cross-gender word recognition. *J. Speech Lang. Hear. Res.* 55:554–60

van Heugten M, Johnson EK. 2014. Learning to contend with accents in infancy: benefits of brief speaker exposure. *J. Exp. Psychol. Gen.* 143:340–50

Vihman MM, McCune L. 1994. When is a word a word? *J. Child Lang.* 21:517–42

Waxman SR, Gelman SA. 2009. Early word-learning entails reference, not merely associations. *Trends Cogn. Sci.* 13:258–63

Weisleder A, Fernald A. 2013. Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychol. Sci.* 24:2143–52

Werker JF, Curtin S. 2005. PRIMIR: a developmental framework of infant speech processing. *Lang. Learn. Dev.* 1:197–234

Werker JF, Tees RC. 1984. Cross-language speech-perception—evidence for perceptual reorganization during the 1st year of life. *Infant Behav. Dev.* 7:49–63

Willits JA, Seidenberg MS, Saffran JR. 2014. Distributional structure in language: contributions to noun–verb difficulty differences in infant word recognition. *Cognition* 132:429–36

Wojcik EH, Saffran JR. 2015. Toddlers encode similarities among novel words from meaningful sentences. *Cognition* 138:10–20

Yang CD. 2004. Universal Grammar, statistics or both? *Trends Cogn. Sci.* 8:451–56
Yeung HH, Werker JF. 2009. Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition* 113:234–43
Yu C, Ballard DH. 2007. A unified model of early word learning: integrating statistical and social cues. *Neurocomputing* 70:2149–65
Yurovsky D, Smith LB, Yu C. 2013. Statistical word learning at scale: The baby's view is better. *Dev. Sci.* 16:959–66