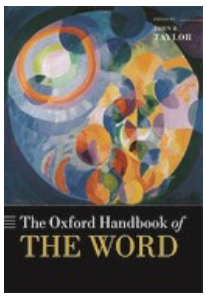

The Oxford Handbook of the Word

John R Taylor (ed.)

https://doi.org/10.1093/oxfordhb/9780199641604.001.0001

Published: 2014 **Online ISBN:** 9780191756696 **Print ISBN:** 9780199641604

CHAPTER

31 How Infants Find Words

Katharine Graf Estes

https://doi.org/10.1093/oxfordhb/9780199641604.013.014 Pages 536–549

Published: 03 March 2014

Abstract

A fundamental skill in early language acquisition is the ability to detect individual words in the fluent stream of speech. This process is challenging because the speech signal does not contain fully reliable acoustic word boundary indicators. Yet before 1 year of age, infants begin to find words and to make sense of those words. There is evidence that infants segment speech by taking advantage of language-general word boundary cues, such as isolated words, utterance edges, and patterns of syllable probabilities, as well as language-specific cues such as lexical stress and phonotactics. Infants use these sources of information to extract new word forms, making them available for further linguistic processing. The present chapter addresses the development of infants' ability to segment words, as well as the connections between word segmentation and vocabulary acquisition.

Keywords: word segmentation, infants, vocabulary acquisition, fluent speech, word boundary cues, phonotactics, lexical stress

Subject: Language Acquisition, Linguistics

Series: Oxford Handbooks

31.1 Introduction

TO learn what a word means, or to learn how a word functions in sentences, infants must first discover individual words in the continuous stream of speech. For adults, detecting words seems like a trivial problem because they can rapidly use stored knowledge of known words to segment the speech signal. However, when hearing an unfamiliar language, adults encounter the same problem that infants face in their native language. It is difficult to tell where one word ends and the next begins because spoken language does not contain fully reliable acoustic cues to word boundaries. The segmentation problem is even more difficult for infants than for adults because infants lack a priori knowledge that the speech stream consists of individual units, that they must detect these units, then decipher their meanings, and link them together in sentences to extract higher-level structure. Yet before 1 year of age infants begin to find words and to make sense of the words they detect. This crucial achievement sets the stage for further advances in language acquisition. The present chapter addresses the development of infants' ability to segment words. It examines the information that infants use to discover words, as well as the connections between word segmentation and vocabulary acquisition.

A seminal study by Jusczyk and Aslin (1995) provided early evidence that infants can detect words in continuous speech. Jusczyk and Aslin tested 6- and 7.5-month-olds' ability to recognize word forms within fluent speech passages when they had previously heard the words presented in isolation. This series of experiments, and many subsequent word segmentation experiments, used the head-turn preference procedure. In the procedure, infants were first familiarized with a pair of native language (English) words presented in isolation, with pauses between repetitions (e.g. *bike ... bike ...*). During testing, infants heard passages that contained the familiarized target words (e.g. 'The girl rode her big *bike*') and passages that contained novel words ('The *cup* was bright and shiny'). Differences in infants' listening times to the
p. 537 passages indicated whether they detected the familiarized words. To measure listening time, the test ↳ passages were played from speakers mounted near blinking lights on the left and right sides of the room. When the infant looked toward a light, a test passage played. When the infant looked away, the passage stopped. Thus, infants controlled the duration of their listening time to the stimuli. Jusczyk and Aslin found that 7.5-month-olds listened longer to the passages with familiarized target words than to passages with novel words. This demonstrates that infants detected the target words within the passages. Six-month-olds did not discriminate the two types of passages, thereby giving no indication that they had detected the target words. Over this brief time window, between 6 months and 7.5 months, infants develop the ability to detect words in continuous speech. Jusczyk and Aslin (1995) also found that 7.5-month-olds could perform the reverse process; when they were first familiarized with words embedded in passages, they listened longer to isolated repetitions of the target words than to novel words. This measure of infants' attention to familiar and novel words has revealed the capabilities and limitations of infant word segmentation. Jusczyk and Aslin's (1995) study provided the foundation for a rich literature exploring the emergence of infants' ability to find words in continuous speech.

31.2 Early use of known words to detect new words

One of the earliest processes that infants use to detect words is quite similar to the primary process that adults use. Given their large lexicons, adults frequently apply top-down knowledge of stored lexical units to segment words, including unfamiliar words (e.g. Marslen-Wilson and Welsh 1978; McClelland and Elman 1986). For known words, boundaries are already well defined. Adults can also readily use the known words surrounding a novel word to identify the new item. Infants lack large vocabularies, but some words are highly frequent in infants' input and may be detected quite early. For example, infants hear many repetitions of their own names and names for their caregivers. Bortfeld et al. (2005) found that 6-month-old infants can recognize their own names as well as the highly familiar word *Mommy* in passages of running speech. Furthermore, they found that infants can take advantage of these early words to detect new words. When infants heard passages containing the familiar word (*Mommy* or the child's own name) followed by a consistent target word (e.g. *Mommy's feet*), they used the familiar word to segment the target word. During testing, infants listened longer to repetitions of target words than to repetitions of novel words. Infants did not show this discrimination pattern when presented with an unfamiliar name preceding the target during familiarization, even if it was phonologically similar to a familiar name (e.g. *Tommy's feet*). The infants required the anchor of a salient, highly familiar word in order to support the discovery of a new word. This study demonstrates that young infants can use stored word knowledge to segment continuous
p. 538 speech. Because ↳ of infants' small vocabularies, their use of known words is limited, but the mechanism is consistent with how adults segment speech.

31.3 Learning from words in isolation

Stored word knowledge supports the ability to learn more words. However, infants cannot fully rely on known words to segment speech. How do infants break into the speech signal when they only know a few words? Words produced in isolation may provide novice language learners with a way to start identifying and storing words. When words occur surrounded by silence, the signal is pre-segmented, allowing infants to learn about the word's phonological form without working to extract it from the surrounding context. Brent and Siskind (2001) have proposed that infants acquire an initial vocabulary of words heard in isolation that they subsequently use to detect other words. In support of this argument, Brent and Siskind analysed samples of mothers speaking to their 9- to 15-month-old infants. They found that about 9 per cent of mothers' speech to infants consisted of words in isolation. Importantly, a substantial number of the words in infants' early vocabularies were words that had previously occurred in isolation. Furthermore, the frequency with which a mother produced a given word in isolation reliably predicted the child's ability to produce that word later in development; sheer frequency of exposure was not a reliable predictor of word use. Lew-Williams, Pelucchi, and Saffran (2011) have also provided experimental evidence that hearing words in isolation can help infants to detect words in fluent speech, even when they are listening to an unfamiliar language.

31.4 Detecting words at utterance edges

In addition to words in isolation, words that occur at the beginnings and ends of utterances provide relatively clear word boundary information. One word boundary is available 'for free' because it follows or precedes a pause. There is evidence that the edges of utterance boundaries are particularly useful for infants' word segmentation. Seidl and Johnson (2006) found that 8-month-old infants segmented words at utterance edges more readily than words that occurred in the middle of sentences. The effect was the same for words that occurred utterance-initially or utterance-finally. Seidl and Johnson explained that these findings support the Edge Hypothesis: infants use utterance edges to detect words in fluent speech, therefore words that occur at utterances edges will be easier to segment than completely embedded words. This may occur because of the silences that precede and follow utterances, or because of the strengthening and word lengthening that commonly occur at the beginnings and ends of utterances. There could also be a broad cognitive or perceptual bias to attend to the beginnings and ends of ↳ sequences. Many of these facilitative effects could also occur at other types of boundary, such as those that occur at mid-utterance pauses or at intonational phrases (e.g. Nazzi et al. 2000). These possible explanations are not mutually exclusive. There may be a combination of effects driving the facilitation of word segmentation at utterance edges. Regardless of the underlying cause, infants' use of edges has potential to be a powerful cue for detecting words.

p. 539

However, words in isolation and at the edges of utterances are insufficient to fully support the development of word segmentation skills. Even infant-directed speech primarily consists of continuous sequences that contain multiple words (Brent and Siskind 2001; see also Fernald and Morikawa 1993). Many of the words that infants must learn will not frequently occur at edges, in isolation, or surrounded by highly familiar words. However, there are numerous cues to word boundaries that provide additional support for detecting words. Although none of these cues is fully reliable, they provide probabilistic word boundary indicators that can be combined to facilitate segmentation.

31.5 Stress cues to word boundaries

One source of information for identifying word boundaries is stress patterns. From very early in life, infants are sensitive to prosody (Mehler et al. 1988; Nazzi, Bertoncini, and Mehler 1998). For example, lexical stress is a salient acoustic characteristic of the speech signal, and infants discriminate words with different lexical stress patterns by 2 months of age (Jusczyk and Thompson 1978). Lexical stress provides a strong segmentation cue. For example, most English bisyllabic words follow a trochaic stress pattern in which strong (stressed) syllables precede weak (unstressed) syllables (e.g. BAby). The opposing iambic pattern, in which weak syllables precede strong syllables, occurs in English but is less common (e.g. guiTAR). English-speaking adults use stressed syllables to identify word onsets (Cutler and Norris 1988; McQueen, Norris, and Cutler 1994), and there is evidence that infants do as well (Curtin, Mintz, and Christiansen 2005; Echols, Crowhurst, and Childers 1997; Morgan 1996). Jusczyk, Houston, and Newsome (1999) found that 7.5-month-olds segmented trochaic syllable sequences from fluent speech passages. When presented with iambic sequences, they mis-segmented, interpreting the stressed syllable as the word-initial syllable. By 10.5 months, infants could correctly segment iambic words, possibly by integrating other cues to word boundaries. However, stress has been found to be a particularly powerful word segmentation cue. At around 8–9 months of age and beyond, infants weigh stress cues more heavily than syllable-level patterns (Johnson and Jusczyk 2001; Johnson and Seidl 2009) or phoneme-level word boundary markers (Mattys et al. 1999). Thus, attention to where stressed syllables fall in the speech signal provides important information regarding where words begin and end.

31.6 Statistical cues to word boundaries

In order for infants to use native language stress patterns to detect words, they must be able to determine what the predominant lexical stress patterns are. The same principle holds for other language-specific cues to word boundaries (discussed in more detail below). Infants must know how cues align with word onsets and offsets in order for the information to be useful. This seems to require that infants extract a vocabulary of stored phonological forms and detect the recurrent patterns across them, such as 'English words frequently have stressed first syllables'. Isolated words and words at utterance edges will provide infants with some segmented items, but others must be extracted from continuous speech. One source of information for establishing an initial lexicon of segmented words is based on patterns of sound co-occurrences in the language. Over speech corpora, sound sequences within words occur together more reliably than sound sequences that cross word boundaries (Harris 1955; Swingley 2005). Moreover, young infants, as well as children and adults, are able to detect these patterns of transitional probability cues to word boundaries (Aslin, Saffran, and Newport 1998; Saffran, Aslin, and Newport 1996; Saffran et al. 1997).

To test infants' ability to use transitional probabilities in word segmentation, Saffran et al. (1996) presented 8-month-olds with an artificial language in which the only reliable word boundary cues were the patterns of transitional probabilities within and across words. Within words, the transitional probability from one syllable to the next was perfect (1.0) because these syllables reliably co-occurred. For example, the infants heard fluent speech sequences like *golatu#pabiku#tibudo#daropi*. Within the word *golatu*, the syllable *go* was always followed by *la* and *tu*. However, the syllable *tu* could be followed by any of the three remaining word-onset syllables in the language (*pa, ti, da*). Therefore, the transitional probability from one syllable to the next was lower (.33) across word boundaries. After only 2 minutes of listening to the language, infants demonstrated learning of its structure. During testing, infants discriminated between words from the language and low transitional probability part-word sequences that had crossed word boundaries (e.g. *bu#pabi*), showing longer listening times to the novel part-words. There is recent evidence that infants as young as 5 months can use transitional probability information in statistical word segmentation tasks (Thiessen and Erikson 2013; and perhaps even younger, see Teinonen et al. 2009).

Infants may use transitional probability information to segment words from continuous speech before they have access to other cues (Saffran and Thiessen 2003). It is not necessary to generalize a pattern across vocabulary items (e.g. stressed syllables usually start words) in order to use transitional probability. Rather, it can act as a language-general mechanism that is available before infants have learned word boundary indicators that are specific to their native languages. Thiessen and Saffran (2003) reported that when transitional probability and stress cues were placed in conflict (i.e. when stress occurred on the second syllable of statistically defined words), ↳ 7-month-olds weighed the transitional probability patterns more heavily than the stress patterns, but 9-month-olds showed the opposite pattern. Johnson and Jusczyk (2001) and Johnson and Seidl (2009) also found that 9- and 11-month-olds followed stress cues, rather than syllable-transitional probability cues. Thiessen and Saffran (2003, 2007; see also Thiessen and Erickson 2013) have proposed that infants use transitional probability information to acquire a small vocabulary of stored word forms. Based on this vocabulary, they can extract regularities, such as how stressed syllables correlate with word boundaries, as well as other language-specific cues that can be used to detect words in fluent speech. As knowledge of language-specific cues strengthens, reliance on transitional probability may decline.

31.7 Phonotactic cues

In addition to syllable-level probabilities, language-specific phoneme patterns also mark words. The phonotactic patterns of a language include the constraints on the allowable locations and combinations in which phonemes can occur in a given language, as well as the frequency of occurrence of phonemes and phoneme combinations. Phonotactic information can serve as a word boundary cue once infants have learned that some phonemes and phoneme combinations do not occur word-initially or word-finally in their native language. For example, in English, words do not start with the phoneme /ŋ/ (the final consonant in *sing*), and cannot end with the consonant /h/ (the initial consonant in *hat*). English words also do not begin or end with certain consonant clusters. This information can identify word boundaries. For example, if a listener encounters a phoneme sequence such as /bd/, phonotactic information indicates that a word boundary is present because words do not begin or end with /bd/ in English. However, the cluster can occur in continuous speech across a word boundary, as in the phrase *bad boy*. Thus, knowledge of the constraints on phoneme combinations and locations can guide listeners to the correct interpretation of word onsets and offsets.

Adults use phonotactic information in word recognition. McQueen (1998) demonstrated that participants were faster to identify words in fluent speech when the sequences contained phonotactic cues to word boundaries. Mattys and Jusczyk (2001) investigated this process in 9-month-olds. Infants listened to passages of speech that contained target (nonsense) words. One word (e.g. *gaffe*) was embedded in a phonotactic context that supported word segmentation. That is, the phoneme sequences in which the word was embedded formed consonant clusters at the target word onsets and offsets that rarely (or never) occur within native-language (English) words. Rather, the phoneme sequences typically occur across word boundaries. For example, the target word *gaffe* occurred in sentences such as *A spun gaffe heads the list of new inventions*. The sequences /n/–/g/ and /f/–/h/ do not typically occur within English words, and so these consonant sequences indicate likely word boundaries. In contrast, the other word (e.g. *tove*) was embedded
p. 542 in a phonotactic context that did not support word segmentation. ↳ Rather, the phoneme sequences at the onset and the offset of the target word typically occurred within English words. The target word *tove* occurred in sentences such as *A gruff tove knows most forest animals*. The consonant clusters /f/–/t/ and /v/–/n/ can occur at word boundaries, but they also occur within English words, allowing the target word to blend into the surrounding words. In testing, infants heard repetitions of the two target words as well as repetitions of novel words. Compared to the novel words, infants listened longer to the words that had been embedded in the phonotactic contexts that supported segmentation. They did not differentiate the novel words from the words embedded in the poor phonotactic segmentation contexts. This suggests that the infants recognized the words from the good phonotactic segmentation contexts when they were subsequently produced in isolation, but not the words from the poor contexts. Thus, by 9 months of age infants are sensitive to language-specific phoneme patterns, and can use them in word segmentation.

31.8 Allophonic cues

Another type of phonemic word boundary cue present in fluent speech is allophonic variation. Phonemes are realized differently depending on word position. For example, the /t/ produced at word onset is aspirated (*tip* is pronounced [t^hɪp]), but the word-internal /t/ in *stop* is not. Jusczyk, Hohne, and Bauman (1999) tested the sensitivity of 9- and 10.5-month olds to these subtle differences in phoneme realizations as they occur in fluent speech (see also Hohne and Jusczyk 1994). The infants were presented with a two-syllable item that could be produced with two different allophonic variations of the phonemes /t/ and /r/: half of the infants heard *night rates* and half heard *nitrates*. They were also familiarized with an unrelated word (e.g. *hamlet*). During testing, infants heard passages that contained both allophonic variation target words as well as the unrelated words. The 10.5-month-olds listened significantly longer to the alternative variation of the target word; that is, infants exposed to *night rates* treated *nitrates* like an unfamiliar word (and vice versa). The 9-month-olds displayed no difference in listening time. This series of experiments shows that by 10.5-months of age, infants are sensitive to the patterns of allophonic variations that occur within and across words in fluent speech. Therefore, they may be able to use this information to detect words.

31.9 Integrating word boundary markers

The studies reviewed above indicate that there are many cues to word boundaries present in the speech stream. However, none of these cues is fully reliable. A given child will not hear every word produced in isolation, some words will have non-dominant stress patterns, and some will begin with infrequent p. 543 phoneme sequences that also occur ↳ across word boundaries. If an infant relies too strongly on any individual cue, she is in danger of mis-segmenting or failing to segment a substantial portion of her input. However, there is power in integrating multiple probabilistic word boundary indicators. One way that multiple cues can be beneficial is that learners may use the presence of one cue to detect new informative cues. For example, as discussed above, Thiessen and Saffran (2003, 2007) have proposed that infants use transitional probability information to learn about lexical stress. Lew-Williams et al. (2011) found that infants can integrate words in isolation with transitional probability information to facilitate word segmentation. In addition, Sahni, Seidenberg, and Saffran (2010) reported that infants could use a language-general segmentation cue, transitional probability, to discover a novel language-specific cue to word onsets.

Studies examining the integration of word segmentation cues have largely addressed how infants interpret cues when they are placed in conflict (e.g. Johnson and Jusczyk 2001; Thiessen and Saffran 2003; Mattys et al. 1999). However, there is evidence that infants can integrate converging information to detect cohesive sequences in speech. Morgan and Saffran (1995; see also Morgan 1994) presented 6- and 9-month-olds with syllable strings that contained correlated cues; both syllable sequence patterns and rhythmic patterns indicated that the syllables formed a coherent unit. In other conditions, infants heard syllable strings in which the cues conflicted or were unavailable. One cue indicated that the syllables should be grouped together, but the other did not, or no cues were present. The 9-month-olds only displayed evidence of grouping the syllable sequences into cohesive units when both cues were present. They seemed to require the availability of the converging cues to word structure. In contrast, 6-month-olds did not take advantage of the conjunction of the patterns as effectively as the older infants. The ability to integrate word segmentation cues may require developmental time and experience to fully emerge. Studies demonstrating how infants integrate convergent information or use an early segmentation strategy to detect new cues support the idea that the complexity present in the ambient language is important for development. Rather than hindering acquisition, the richness of the linguistic signal may facilitate word segmentation, and language acquisition more broadly (Sahni et al. 2010; Smith and Yu 2008; Yu and Smith 2007).

Computational models also indicate that multiple converging information sources can facilitate word segmentation. For example, Christiansen, Allen, and Seidenberg (1998) designed a connectionist model of word segmentation that was able to reliably identify word boundaries based on a conjunction of cues. The model, a simple recurrent network, was exposed to a corpus of infant-directed speech. The input included information about phonemes, lexical stress patterns, and utterance boundaries. When all sources of information were available, the network could reliably identify word boundaries. However, none of these sources of information alone was sufficient to mark the words. Perruchet and Tillmann (2010) also examined the abilities of computational models and human adults to integrate word boundary cues such as transitional probabilities, the perceived word-likeness of sequences, and the context that emerged when p. 544 word boundary markers were present for some but not all words. Human participants ↳ and the models were better able to detect word-like chunks when multiple cues to word boundary locations were available than when only transitional probability information supported segmentation.

Many computational models have addressed the word segmentation problem (reviewed in Brent 1999). The models differ in characteristics such as the nature of the input, learning algorithms, and how the output of learning is represented (e.g. Frank et al. 2010; Mirman et al. 2010; Monaghan and Christiansen 2010; Perruchet and Tillmann 2010; Räsänen 2011; Rytting, Brew, and Fosler-Lussier 2010). But a common theme in many models is the assumption that in order to process speech, it is necessary to identify individual words in the speech stream. Moreover, young learners must detect words in order to associate meanings with them and to determine how they function in sentences. Swingley's (2005) computational model focused on the effectiveness of infants' word segmentation mechanisms for extracting real words from infant-directed speech. Swingley proposed that infants use clustering strategies to detect coherent sound sequences. The model indicated that regularities present in syllable co-occurrences can support the detection of a substantial number of real words. Swingley proposed that infants use patterns of syllable probabilities to detect statistically cohesive sound sequences that increase in familiarity. The items enter a 'protolexicon' of words that have a representational foundation in memory (Swingley 2005: 118). These stored units may then be available for further linguistic processing, such as being associated with meanings or in tracking syntactic patterns.

31.10 Learning meanings for segmented words

Experimental evidence further strengthens the notion that infants detect and store word forms and then apply those representations to the process of associating forms with meanings during word learning. Swingley (2007) found that 19-month-olds showed stronger learning of object names when the infants first heard the names embedded in speech passages before they acted as object labels. Exposure to the phonological forms led to more precise representations of the labels: infants detected mispronunciations that they failed to detect without prior experience hearing the words in passages. Several recent experiments have probed the connection between word segmentation and word learning by investigating how infants use specific word segmentation cues, such as syllable transitional probabilities, phonotactics, or prosody, to extract words and then apply those units to support linguistic functions.

In one such experiment, Graf Estes et al. (2007) examined infants' ability to detect words using statistical cues, then to use those words as object labels. The 17-month-old infants first participated in a word segmentation task. They listened to an artificial language in which the only cue to word boundaries was the p. 545 pattern of syllable transitional ↳ probabilities within words (high probability) vs. across-word boundaries (low probability), similar to the statistical learning experiments described previously (e.g. Saffran et al. 1996). After listening to the speech stream, they immediately participated in an object-labelling task. For one group of infants, the objects were labelled with words from the artificial language. For the other group of infants, the objects were labelled with low transitional probability sequences that crossed word

boundaries or syllable sequences that never occurred together in the speech stream. Graf Estes et al. (2007; see also Graf Estes 2012) found that infants only learned the labels when they had prior opportunity to segment them from the speech stream; they learned the high-probability words as labels, but not the low-probability sequences. Hay et al. (2011) found a similar pattern of results when infants listened to an unfamiliar natural language (English-learning infants presented with Italian). Graf Estes and colleagues have proposed that by tracking statistical regularities, infants segment coherent units that are ready to be associated with meanings.

Graf Estes (2014) also recently found that infants can use native language phonotactic patterns to detect words in fluent speech, and then use the word form representations to facilitate label learning. The infants (14-month-olds) listened to passages in which target words were surrounded by words that produced good phonotactic word boundary contexts or poor word boundary contexts, based on the passages from Mattys and Jusczyk's (2001) original study of phonotactic word segmentation. When infants heard the target words embedded in good word boundary contexts, they learned the words as object labels. When the words were embedded in contexts that did not support segmentation, they failed to learn the labels. They also failed to learn the labels in the absence of any prior exposure to the labels. The results of this experiment indicate that novice word learners, such as 14-month-olds, are able to take advantage of language-specific word segmentation cues, then store those representations until a potential referent is available.

Recently, Shukla, White, and Aslin (2011) presented evidence that even younger infants, 6-month-olds, can use word segmentation cues to detect a new word and associate it with an object. The infants listened to short utterances in an artificial language while watching animated sequences that contained a target object and two unlabelled distractor objects. The target object reliably co-occurred with a bisyllabic sequence that had perfect internal transitional probability (i.e. a word), but was surrounded by syllables that varied. The target object moved during labelling while the distractor objects remained still. During testing, the infants viewed the target and a distractor object and heard the reliable syllable sequence or a syllable sequence that had crossed word boundaries in the artificial language. When they heard the high transitional probability word sequence, infants increased their looking to the target object. They did not increase looking to the target when they heard the across-word sequence. However, this pattern of results only occurred for infants who, during training, heard the word at the end of a prosodic pattern that indicated the end of a phrase. When the word crossed a prosodic phrase boundary during training, the infants did not associate it with the target object (rather, they seemed to map the word to the distractor; see Shukla et al. 2011 for ↳ additional details). The prosodic pattern marking the edges of phrases seemed to be crucial for associating the word form and referent. Shukla et al.'s study indicates that even very young infants have the potential to segment linguistic units and associate them with referents. Infants who have only just started to comprehend a few common and salient native language words (Bergelson and Swingley 2012; Tincoff and Jusczyk 1999, 2012) already possess mechanisms for segmenting words from fluent speech and learning the roles that those words play.

p. 546

Despite the strength of infants' word segmentation abilities, they do not always segment the speech stream correctly. Infants may mis-segment when a word is inconsistent with their expectations about word form characteristics. For example, Jusczyk et al. (1999) reported that infants seemed to segment the sequence *taris* from the phrase *guitar is* because the dominant English stress pattern indicates that the second (stressed) syllable in *guitar* is actually the onset of a word (see also Cutler and Butterfield 1992, for evidence with adults). In other experiments, infants fail to detect words when limited segmentation cues are available (e.g. Johnson and Tyler 2010; Mattys and Jusczyk 2001). There are also classic examples of children's segmentation errors that appear in speech production. A child who is instructed to behave replies, 'I am /heɪv/' (Peters 1983: 43). Children sometimes treat common phrases such as *look at*, *want to*, and *do it* as single words before eventually decomposing them into their parts (Peters 1985; see also Brown 1973). Such segmentation errors do not occur randomly. Rather, they represent sensible generalizations of word

segmentation processes. In the first example above, experience with phrases in which *be* precedes adjectives, such as *be careful* and *be nice*, leads children to parse *behave* as 'be + have' /heɪv/, in which *have* is a new adjective. In other cases, highly frequent exposure to common phrases leads children to interpret them as cohesive units. For example, children may rarely hear *look* when it is not followed by *at*. The input patterns indicate that *look at* is a lexical unit, and children treat it accordingly. Infants and children must accumulate greater experience with the speech signal in order to override these early errors. Given the complexity of the speech signal and the robustness of many word segmentation cues, it is remarkable that children do not display these kinds of mistakes more frequently. Such segmentation errors are notable for what they reveal about how children construct their knowledge of the speech signal.

31.11 Word segmentation and vocabulary development

The studies reviewed above establish that infants possess many tools for finding words in fluent speech. There is also evidence of a close connection between word segmentation and word learning. Infants across a range of ages show superior learning of object labels when they have had prior opportunity to segment those labels from fluent speech (Graf Estes et al. 2007; Hay et al. 2011; Shukla et al. 2011; Swingley 2007).
p. 547 Word ↳ segmentation may provide a foundation for vocabulary acquisition. Infants may extract a store of word forms that are not yet associated with meanings. When the opportunity arises, because of increases in experience or cognitive skills, those previously stored word forms could then readily be available to associate with meanings. This prior store of word forms may contribute to why vocabulary development accelerates so dramatically for many children during the second year of life. Infants do not start word learning as blank slates at around 1 year of age. Rather, they have accrued months of experience listening to speech. This exposure leads to the discovery of many individual word forms, as well as to expectations about the phonological and prosodic forms that words are likely to take, thereby facilitating learning of words that are consistent with these patterns (Graf Estes, Edwards, and Saffran 2011).

Word segmentation is a fundamental process in language acquisition. It follows that word segmentation ability should be associated with language acquisition progress. Infants who have difficulty segmenting words from fluent speech may be at a disadvantage in vocabulary development because they do not have stored word forms available to associate with meanings. In addition, they may have difficulty during online sentence processing if they are slow to identify the speaker's referent. Infants who are skilled at word segmentation will have greater opportunities to learn about the meaning, context, and grammatical roles of the words they detect than infants who are not (see a related argument in Fernald, Perfors, and Marchman 2006; Marchman and Fernald 2008).

Newman et al. (2006) reported that children's language abilities can be linked to their segmentation skills in infancy. Infants participated in word segmentation tasks that tested the use of cues such as prosody or phonotactics, or the ability to recognize words across changes in the speaker's gender. In this retrospective study, infants were initially tested between 7.5 and 12 months. At age 2, infants who had larger vocabularies (the top 15 per cent of the sample) were more likely to have shown successful segmentation as infants (as indicated by following the novelty or familiarity preference pattern that was reliably exhibited at the group level) than infants with smaller vocabularies (the bottom 15 per cent of the sample). Children at the high and low ends of vocabulary size differed in their ability to segment words from fluent speech as younger infants.

In a subsequent study, Newman and colleagues (2006) re-tested a subset of the same group of children between 4 and 6 years of age. To examine whether associations between segmentation performance and language skills could be due to general cognitive differences between children who were identified as 'segmenters' versus 'non-segmenters' as infants, the children were tested on measures of linguistic and non-linguistic cognitive abilities. Children who were segmenters as infants had higher scores on languages

measures than the non-segmenters, but the two groups did not differ in general intellectual abilities, or non-verbal intelligence. The segmenters had higher scores for both vocabulary and syntax. Thus, Newman et al.'s research supports the idea that early segmentation performance lays a foundation for subsequent language development. Children who have difficulty identifying words in fluent speech may then find it challenging to associate meanings with new words, to learn their grammatical roles, and to combine them with other morphemes to form complete utterances.

p. 548 Singh, Reznick, and Xuehua (2012) recently reported a longitudinal study examining the association between word segmentation at 7.5 months and vocabulary size at 24 months. Infants were familiarized with a pair of words presented in isolation, then tested on their ability to recognize the words embedded in passages. Infants also participated in a second, more challenging task. They were familiarized with a pair of words, but in testing, one of the words was produced in a different pitch, and the other word was produced in the same pitch as during familiarization. To succeed, infants had to recognize the familiarized word and generalize their representation of the word across acoustic variation. Singh and colleagues found that infants' ability to recognize the target words at 7.5 months (measured as the looking time difference to target words vs. unfamiliar words) was significantly correlated with productive vocabulary size at 24 months. For the simple segmentation task in which the speaker's voice was consistent from familiarization to test, performance was also correlated with general cognitive skills. However, for the complex task that required generalization across voices, performance was associated with vocabulary size, but not general cognitive ability.

The work of Newman and colleagues and Singh and colleagues provides further evidence that early language processing abilities are meaningfully related to later language acquisition progress. Previous studies have demonstrated similar associations in a variety of tasks: the ability to process rapid auditory transitions (Benasich and Tallal 2002), native and non-native phoneme discrimination (e.g. Kuhl et al. 2005; Tsao, Liu, and Kuhl 2004), and speed of lexical access (Fernald and Marchman 2012; Fernald et al. 2006; Marchman and Fernald 2008) in infancy are all associated with later language outcomes. Segmentation ability follows the same pattern. Future research will be necessary to investigate whether infants' attention to particular word segmentation cues (e.g. transitional probability, lexical stress) are also associated with later language skills. The findings from these correlational studies indicate that measures of fundamental language processing abilities, such as word segmentation, have the potential to be extended for practical application. By understanding individual variation in such skills, it may be possible to gain greater understanding of development in children at risk of lasting language impairments.

31.12 Conclusion

For adults, the act of segmenting speech usually occurs automatically. Considering word segmentation from the infant's perspective emphasizes the magnitude of the challenge that word segmentation poses for language processing and acquisition. Despite the complexity of the speech signal and the lack of obvious word boundary markers, infants discover individual words. Before age 1, they are sensitive to many cues to where words begin and end. Infants detect and use language-general cues, such as isolated words, utterance edges, and transitional probabilities. They also learn and apply language-specific cues, such as lexical stress p. 549 and phonotactics. Although no cue ↳ is fully reliable, their convergence is highly informative. Infants' precocious learning mechanisms and speech processing capabilities allow them to access these information sources to pull out individual words from continuous speech—words that can then be associated with meanings and added to the lexicon. Thus, the ability to find words is an essential, fundamental skill in language acquisition.