Journal of Experimental Child Psychology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Brief Report

Nine-month-olds use frequency of onset clusters to segment novel words

Stephanie L. Archer [a,*], Suzanne Curtin [b]

[a] Department of Psychology, University of Warwick, Coventry CV4 7AL, UK
[b] Department of Psychology, University of Calgary, Calgary, Alberta T2N 1N4, Canada

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Language development
Infant speech perception
Word segmentation
Frequency
Phonology
Phonotactics

ABSTRACT

Before their first birthday, infants have started to identify and use information about their native language, such as frequent words, transitional probabilities, and co-occurrence of segments (phonotactics), to identify viable word boundaries. These cues can then be used to segment new words from running speech. We explored whether infants are capable of detecting a novel word form using the frequency of occurrence of the onset alone to further characterize the role of phonotactics in speech segmentation. Experiment 1 shows that English-learning 9-month-olds can successfully segment a word from natural speech if the onset is legal in English (i.e., *pleet*) but not if the onset is illegal (i.e., *tleet*). Experiment 2 shows that English-learning 9-month-olds are successful at word segmentation when presented with two onset clusters that vary in statistical frequency. Infants familiarized to a high-frequency onset (i.e., *trom*) were successful at segmenting the target word embedded in speech, but those familiarized to the low-frequency onset (i.e., *drom*) were unsuccessful. Together, these results show that infants use statistical information from the speech input and that low levels of exposure to onset phonotactics alone might not be sufficient in identifying word boundaries.

* Corresponding author.
E-mail address: s.archer@warwick.ac.uk (S.L. Archer).

http://dx.doi.org/10.1016/j.jecp.2016.04.004
0022-0965/

Please cite this article in press as: Archer, S. L., & Curtin, S. Nine-month-olds use frequency of onset clusters to segment novel words. *Journal of Experimental Child Psychology* (2016), http://dx.doi.org/10.1016/j.jecp.2016.04.004

Introduction

Adults have large dynamic lexicons that bolster their knowledge of the language-specific cues that are necessary for speech processing. Cues like phonotactics[1] and prosody help adults to detect words in continuous speech, especially when interacting with probabilistic information (Finn & Hudson Kam, 2008; Shukla, Nespor, & Mehler, 2007). Infants, however, experience a vast mixture of acoustic information and are in the process of determining what "word finding" cues are available in their native language. They must rely on different sources of information to discover cues and patterns that aid in word segmentation (for a review, see Curtin & Archer, 2015). Using highly frequent words (e.g., *mommy;* Bortfeld, Morgan, Michnick Golinkoff, & Rathbun, 2005) and words at utterance edges (Johnson, Seidl, & Tyler, 2014), infants can segment new words from speech at 6 months of age. At around 7–9 months, infants are capable of tracking the predominant stress pattern of their language (Curtin, Mintz, & Christiansen, 2005) and syllable co-occurrences (transitional probabilities [TPs]; Saffran, Aslin, & Newport, 1996) in artificial languages, but when English-learning infants are exposed to TPs in a novel natural language (i.e., Italian), they segment only when isolated words are included in the speech stream (Lew-Williams, Pelucchi, & Saffran, 2011). Previous exposure to phonological templates also helps to inform infants' segmentation (Saffran & Thiessen, 2003). When presented with repeated words within passages of natural speech, infants can extract CVC (consonant–vowel–consonant) target words (e.g., *feet*, *cup*). Furthermore, they recognize only those test items that are an exact match to those with which they were familiarized (Jusczyk & Aslin, 1995).

Thus, by the second half of their first year, infants have experience with and use a number of segmentation cues. Here we focused on infants' use of native-language phonotactic information to identify new words. Specifically, we asked whether infants' knowledge of stop–liquid[2] onsets[3] influences their segmentation of novel words from speech.

Infants show preferences for native-language phonotactics by approximately 9 months of age (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993). They are sensitive to onset cluster frequency (Archer & Curtin, 2011) and show preferences for onset and coda clusters that conform to native phonotactics over those that do not (e.g., *bref* over illegal *febr;* Friederici & Wessels, 1993). They also prefer novel words with high-probability segment co-occurrences to low-probability ones (Jusczyk, Luce, & Charles-Luce, 1994). Bilingual infants also prefer the phonotactic patterns of the dominant language in their input (Bosch & Sebastián-Gallés, 2001).

Phonotactic sensitivity helps infants to detect viable word boundaries in speech. Nine-month-olds detect the difference between consonants that cross a syllable boundary and those that cross a word boundary (Mattys, Jusczyk, Luce, & Morgan, 1999). For example, in English the string CVŋ.**k**VC is an allowable within-word consonant sequence between syllables. However, a string such as CV**f**.**h**VC is only observed in English between words. Infants of 9 months prefer within-word phonotactic combinations, suggesting knowledge of legal word forms and syllable boundary phonotactics (Mattys et al., 1999). Similarly, 9-month-olds exposed to strong phonotactic boundary cues (e.g., *bean**gaffe**hold*) and then tested on isolated target words (e.g., *gaffe*) demonstrate a preference for target words, whereas infants exposed to poor boundary cues (e.g., *fang**gaffe**tine*) do not (Mattys & Jusczyk, 2001). Moreover, by 8 months infants will use high (but not low) phonotactic probability between syllable sequences as viable labels for categories (Erickson, Thiessen, & Graf-Estes, 2014). Together, these studies demonstrate the usefulness of phonotactic information for infants in identifying words and forming categories.

Native-language sensitivity of co-occurrences also informs word mapping and recognition. Twelve-month-olds map novel words containing legal onsets onto objects (e.g., *plot*) but not those containing illegal ones (e.g., *ptak*; MacKenzie, Curtin, & Graham, 2012). After familiarization to an artificial continuous stream of speech, infants of 17 months map highly probable phonotactic labels to objects but not low-probability ones (Graf-Estes, Evans, Alibali, & Saffran, 2007), suggesting that

[1] Phonotactics are language-specific combinations and positions of speech sounds within a word.

[2] Stops are a category of consonants created by stopping airflow and releasing (e.g., *p*, *b*, *t*, *d*, *k*, *g*). Liquids are a category of resonant consonants created by voicing and semi-restricted airflow around the tongue (e.g., *l*, *r*).

[3] Onsets are consonants that are positioned at the beginning of a syllable.

Please cite this article in press as: Archer, S. L., & Curtin, S. Nine-month-olds use frequency of onset clusters to segment novel words. *Journal of Experimental Child Psychology* (2016), http://dx.doi.org/10.1016/j.jecp.2016.04.004