Brain & Language 170 (2017) 39–49

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Infant word segmentation recruits the cerebral network of phonological short-term memory

Yasuyo Minagawa [a,b,c,*], Yoko Hakuno [b], Ai Kobayashi [b], Nozomi Naoi [b,c], Shozo Kojima [a]

[a] Department of Psychology, Faculty of Letters, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa-ken 223-8521, Japan
[b] Graduate School of Human Relations, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
[c] CREST, Japan Science and Technology Agency, Japan

ARTICLE INFO

Article history:
Received 27 August 2016
Revised 1 March 2017
Accepted 2 March 2017
Available online 10 April 2017

Keywords:
Word segmentation
Supra marginal gyrus
Inferior frontal gyrus
NIRS
Phonological short-term memory
Dorsal pathway

ABSTRACT

Segmenting word units from running speech is a fundamental skill infants must develop in order to acquire language. Despite ample behavioral evidence of this skill, its neurocognitive basis remains unclear. Using behavioral testing and functional near-infrared spectroscopy, we aimed to uncover the neurocognitive substrates of word segmentation and its development. Of three age-groups of Japanese infants (5–6, 7–8, and 9–10 months of age), the two older age-groups showed significantly larger temporo-parietal (particularly supramarginal gyrus) responses to target words repeatedly presented for training, than to control words. After the training, they also exhibited stronger inferior frontal responses to target words embedded in sentences. These findings suggest that word segmentation largely involves a cerebral circuit of phonological (phonetic) short-term memory. The dorsal pathway involved in encoding and decoding phonological representation may start to function stably at around 7 months of age to facilitate the growth of the infant's vocabulary.

1. Introduction

Unlike written language, conversational speech does not have noticeable spaces between words. Indeed, speech hardly has any salient acoustic cues to indicate word boundaries. This becomes an obstacle for infants in identifying a spoken word unit. Similar to listening to an unknown language for adults, the native language that infants hear probably sounds like a running sound stream. Hence, a primary step of word acquisition in infants is the ability to segment words from continuous speech. This type of language skill, known as "word segmentation", is one of the significant landmarks in language development in the first year of life.

How does an infant determine where one word ends and another starts? Extensive studies have been performed on this issue, chiefly using the head-turn preference procedure (HPP). Jusczyk and colleagues performed a series of studies and identified several factors that guide infants to segment words such as phonotactic rules, coarticulation, and stress (Jusczyk & Aslin, 1995; Jusczyk, Hohne, & Bauman, 1999; Jusczyk, Houston, & Newsome, 1999). Among them, stress appears to play a significant role in many languages. For example, Jusczyk, Houston, et al. (1999) showed that 7.5-month-old infants were able to identify familiarized English words with a Strong–Weak pattern but not those with a Weak–Strong pattern. Infants whose native language is Dutch show a similar tendency at 9 months of age (Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000). However, languages without any stress accents appear to have different developmental word segmentation processes. French contains no lexical accents, and provides no clear acoustic indications to infants. These results in infants being unable to segment a final syllable by 8 months of age, or find the boundary of a whole word by 16 months of age (Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006).

In the case of Japanese, whose primary acoustic cue to a lexical accent is a pitch pattern, infants start to segment Japanese words with particular phonemic and accentual context around 8–10 months of age (Kajikawa & Masataka, 2003; Sato, Kajikawa, Sakamoto, & Matsumoto, 2007). Specifically, words starting with a High (H) pitch followed by a Low (L) pitch (HL, initial accented pattern) are more easily segmented, whereas those with a Low–High pattern (no initial accentuation) are more difficult to segment. Note that in standard Japanese, a mora with H pitch following a decrement of pitch (i.e. a mora with L pitch) represents an accented mora (e.g. HLL initial accent, LHL second mora accented, LHH no accentuation when the next word starts with H pitch). Vowel qualities of the first and second morae also contribute to saliency of these two successive initial morae, such that an

* Corresponding author at: Department of Psychology, Faculty of Letters, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa-ken 223-8521, Japan.
E-mail address: myasuyo@bea.hi-ho.ne.jp (Y. Minagawa).

http://dx.doi.org/10.1016/j.bandl.2017.03.005
0093-934X/

open-closed vowel sequence is easier to detect than any other sequence. By analyzing the acoustic features of words that are well segmented by 9 months, Sato et al. (2007) concluded that acoustic differences in amplitude and duration between the first and second morae provide a strong cue for infants in finding a boundary. Altogether, Japanese 9-month-old infants have a tendency to segment words under an initial HL accent condition and an open-closed vowel sequence with enough acoustic differences between both morae (Sato et al., 2007). These studies imply that similar to English-learning infants, Japanese infants make use of the acoustic sequence of "Strong–Weak" as a salient cue to initially identify words. Additionally, since Japanese is an agglutinative language, it has case marking auxiliary words, such as topic and subject markers. It has been suggested that these markers also contribute to word segmentation (Kajikawa & Haryu, 2007). Finally, subsequent research after a series of studies by Jusczyk and colleagues showed that depending on the type of stimulus presentation and training method, young infants can segment words earlier (Bergmann & Cristia, 2015; Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Nishibayashi, Goyet, & Nazzi, 2015; Seidl & Johnson, 2006). For example, word–object pairing training with real social interaction enhances word learning, and results in the capability to segment trained words, even for 5–6-month-old Japanese infants (Hakuno, Omori, Yamamoto, & Minagawa, 2017). Here, to compare the accumulating results of previous studies, we focus on a typical word-segmentation task that only uses auditory cues.

Apart from the abovementioned behavioral studies, Kooijman, Hagoort, and Cutler (2005) attempted to investigate the neural basis of word segmentation in infants using event-related potentials (ERP). Employing a similar paradigm as the one by Jusczyk and Aslin (1995), they found a differential electrophysiological signature between trained and control unfamiliar words. During the training phase, positive electrophysiological responses gradually decreased as infants were familiarized with target words, a neural indication of habituation to the word. Accordingly, in the test phase, learned words elicited more negative deflections (around 350–500 ms) in Dutch 7- and 10-month-old infants (Kooijman et al., 2005). HPP shows the earliest evidence for segmentation in Dutch infants is 9 months of age; therefore, ERP is a more sensitive methodology to examine word segmentation. Additionally, Goyet, de Schonen, and Nazzi (2010) tested French 12-month-old infants with no behavioral indication of bisyllabic word segmentation in French (Nazzi et al., 2006), and they showed negative deflections around 350–500 ms in response to familiarized words. Given that these ERP responses are similar to the neural correlate of segmentation observed in the study of Kooijman et al. (2005), French infants were also shown to segment words 4 months earlier by electrophysiological measures than with behavioral measures.

Although these studies crucially reveal the cerebral correlate of word segmentation from continuous speech, ERP cannot precisely detect the brain region engaged in specific cognitive processing. Consequently, the brain mechanism or network underlying word segmentation still remains unclear. Further, it is possible that a brain signature is more sensitive than behavioral results, thus a much younger infants' brain (e.g., 5–6 months of age) may already be responsive to familiarized words inserted in sentences. fNIRS studies of word familiarization strongly support this view, suggesting that neonates exhibit some form of sound memory by presenting isolated words (Benavides-Varela, Hochmann, Macagno, Nespor, & Mehler, 2012). To explore these issues in the present study, we investigated cerebral activation in infants during word training and word segmentation of learned words using both a behavioral method (described in Experiment 1) and a neuroimaging method, specifically, functional near-infrared spectroscopy (fNIRS) (described in Experiment 2). We aimed to reveal: (1) at what age Japanese infants start to segment whole words, as assessed by hemodynamic brain responses and behavioral testing; and (2) what brain network underlies the process of word learning and segmentation, and thus what cognitive processes are engaged within the network.

2. Experiment 1: Materials and methods

In experiment 1, we examined the behavioral performance of an infant's ability to segment whole words from continuous speech. The targeted stimulus words were Japanese three mora words with HLL accents, which are easier to segment (as already explained). Accordingly, we used a forced-choice preferential looking (FPL) paradigm (McCleery, Allman, Carver, & Dobkins, 2007; Teller, 1979) to test Japanese infants from three age groups (5–6, 7–8, and 9–10 months of age).

2.1. Participants

Fifty-four Japanese infants from three age groups (5–6 months: $N = 18$, 14 boys, mean age = 181 days, SD = 22.7; 7–8 months: $N = 18$, 10 boys, mean age = 238 days, SD = 15.2; and 9–10 months: $N = 18$, 13 boys, mean age = 288 days, SD = 14.7) were included in the final data set. They had no hearing or developmental problems, nor significant exposure to foreign languages. There were 21 additional infants, but 15 were excluded because they did not reach the criterion for minimum looking time (2 s) on all 12 test trials. Other exclusions were due to fussiness (5 infants) or insufficient exposure to Japanese at home (1 infant). Parents signed an informed consent form approved by the ethics committee of Keio University, Faculty of Letters [No. 09049].

2.2. Stimuli

The stimulus words used as target words were "*tanishi*" (mud snail or Pilidae) and "*zakuro*" (pomegranate). On the basis of our preliminary study and a previous study (Sato et al., 2007), we used the following two criteria to select these stimulus words. First, our stimulus words were initially accented 3-mora (sub-syllabic unit) nouns with the vowel sequences "a-i" or "a-u" in the first and second morae. Second, the stimulus words had low familiarity scores ranging from 5.0 to 5.5 points, according to the database for spoken words created by Amano and Kondo (1999). The first criterion was chosen on the basis of previous studies, and facilitates infant word identification from sentences. As stated above, Sato et al. (2007) reported that an initial word sequence of "a-u" is easier to detect than "u-a" due to the strong amplitude at the initial mora. This is consistent with our pilot study, where we found a tendency towards easier segmentation of broad–narrow vowel sequences than closed–open vowels in the context of an initially accented word. The second criterion was chosen to ensure that stimulus words were not already familiar to the participants.

For the training session, one of the target words (zakuro or tanishi) was used for each participant. In the test session, the previously familiarized words (zakuro or tanishi) were used as target words, as well as unfamiliarized control words, specifically, "*gaika*" (foreign money) and "*aruji*" (master). These control words adhered to the same criteria for word context as the target words. The combination of familiarized and unfamiliarized control words and frequency of combination use were counterbalanced for each age group as much as possible. Specifically, either "*gaika*" and "*aruji*" was used as the control word for half of the infants in each age group. For test session stimuli, each stimulus word was embedded in six different sentences (Table 1), resulting in a total of 24 sentences. For the stimulus recording, three female native Japanese