

Research report

Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech

Valesca Kooijman^{a,b,*}, Peter Hagoort^a, Anne Cutler^b

^a*F.C. Donders Centre for Cognitive Neuroimaging, Radboud University of Nijmegen, P.O. Box 9101, NL-6500 AH Nijmegen, The Netherlands*

^b*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Accepted 23 December 2004

Available online 11 February 2005

Abstract

Children begin to talk at about age one. The vocabulary they need to do so must be built on perceptual evidence and, indeed, infants begin to recognize spoken words long before they talk. Most of the utterances infants hear, however, are continuous, without pauses between words, so constructing a vocabulary requires them to decompose continuous speech in order to extract the individual words. Here, we present electrophysiological evidence that 10-month-old infants recognize two-syllable words they have previously heard only in isolation when these words are presented anew in continuous speech. Moreover, they only need roughly the first syllable of the word to begin doing this. Thus, prelinguistic infants command a highly efficient procedure for segmentation and recognition of spoken words in the absence of an existing vocabulary, allowing them to tackle effectively the problem of bootstrapping a lexicon out of the highly variable, continuous speech signals in their environment.

© 2005 Elsevier B.V. All rights reserved.

Theme: Neural basis of behavior

Topic: Cognition

Keywords: Word segmentation; Continuous speech; Infants; ERP

1. Introduction

Learning a language from birth entails many steps. One essential step is building a vocabulary of the words of the mother tongue. From the fact that children begin their attempts to talk at around age one, it is clear that the initial steps in vocabulary building have been taken in the first year of life. This is a formidable achievement, especially given the fact that most of the utterances infants hear in the first year of life are not words in isolation, but continuous speech without pauses between the words.

The continuity of speech presents one of the greatest challenges to listeners of all ages and all languages.

Boundaries between individual words in an utterance are not marked by reliable and consistent signals; yet recognizing the individual words which make up an utterance is necessary if the utterance is to be understood. Thus, the individual words must be extracted from the utterance. Fig. 1 illustrates how hard this can be. The three spectrograms in the upper part of the figure represent three isolated utterances of the same word (*hofnar* 'court jester'). The three utterances are not at all the same—they differ both in duration and in spectral quality. The same word also occurs within the sentence which is shown in the lower part of the figure. There are no pauses before or after *hofnar* in the sentence context, and the acoustic shape of the word's onset and offset have been influenced by the preceding and following phonemes.

If it is challenge enough for the adult listener, the continuity of speech presents a very serious problem indeed to the infant listener attempting to build up an initial stock of word forms based on the available input. Word forms must

* Corresponding author. F.C. Donders Centre for Cognitive Neuroimaging, Radboud University of Nijmegen, P.O. Box 9101, NL-6500 AH Nijmegen, The Netherlands. Fax: +31 24 36 10 989.

E-mail address: valesca.kooijman@fcdonders.ru.nl (V. Kooijman).

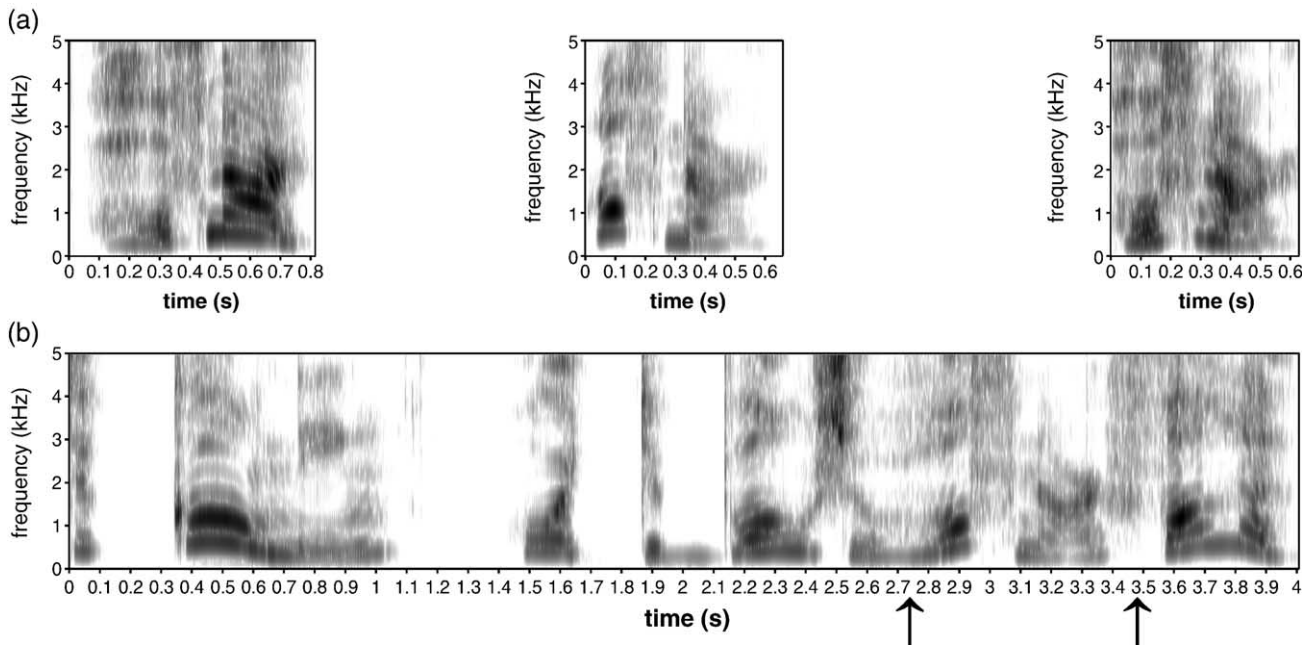


Fig. 1. (a) Spectrograms of three utterances of the word *hofnar* ('court jester') spoken in isolation in an animated infant-directed manner; (b) Spectrogram of the sentence *De koning hoort de boze hofnar vallen* ('The king hears the nasty court jester falling'), spoken in the same style. The displays represent frequency on the vertical axis against time on the horizontal axis, with greater energy represented by darker color. It can be seen from (a) that these three utterances differ considerably, both in duration and in distribution of energy across the frequency spectrum. It can be seen from (b) that most individual words in the sentence adjoin to one another continuously, without a break; the word *hofnar* begins just after 2.7 on the time line and ends just before 3.5. The band of dark energy in the low frequency region (0–50 Hz) coincident with the initial consonant of *hofnar* in the sentence (but absent from the tokens spoken in isolation) represents voicing from the second vowel of *boze* which has continued into the following consonant. Thus, the adjacent phonetic context not only abuts to but also directly affects the form of a word in a sentence.

be recognized as such even though they vary in acoustic form in different contexts, and even though their boundaries in a sentence context are often unmarked. Speech to infants is in this respect not different from speech between adults; in the largest available sample of speech input to an infant listener [16], continuous speech was found to account for 67% of all utterances. Of all the words the infant heard, only 9% of them were uttered in isolation. Thus, the utterance in Fig. 1 – which, as it happens, is taken from the materials of the present study – is a fair approximation of the kind of continuity problem presented daily to infant listeners. (Note that it was thus spoken in an animated, hyper-articulated style characteristic of speech to infants; variability and contextual influence in speech can in fact be far more extreme than is illustrated here).

Nonetheless, infants contrive to cope with this problem, that is, to recognize recurring word forms within continuous speech and to construct an initial set of words which, around the end of their first year, they begin to utter. That is, infants are indeed capable of segmenting words from surrounding speech context. This step in language acquisition is taken in the first year of life, before meaning is attached to words [8]. In this first year, infants start to learn how to segment the continuous speech into discrete units roughly corresponding to individual words. The first indications of word segmentation from context are simply based on acoustic form. There is abundant evidence of

young infants' competence in segmenting and recognizing words, coming principally from studies using the Headturn Preference Paradigm (HPP). This method compares summed listening time for stimuli of one type versus another, with longer listening time taken to indicate a preference. In a two-stage Familiarization and Test version of HPP, infants from 7.5 to 12 months of age have been shown to listen longer to short passages containing words they had just been familiarized with than to similar passages containing unfamiliar words [7–10,12]. This suggests that the infants not only showed a preference for familiar words (over novel words), but also had been able to recognize these newly familiar words even though they were embedded in continuous speech; thus, they must have been able to segment the words from the surrounding continuous speech.

HPP, however, is an indirect measure of segmentation, and it is not possible to investigate with HPP how rapidly segmentation occurs. We wished to look more closely at the time course of word segmentation from continuous speech, and in order to achieve the high temporal resolution necessary for this question, we turned to event-related brain potentials (ERPs). Using ERPs enables us to see what happens in the infant's brain as a particular word in the speech stream is heard; thus, it gives us the opportunity to assess the time needed to segment and recognize this word from speech, as well as to determine whether words are

Table 1
Example of an experimental block of the Dutch materials, with literal English translation between brackets

<i>Familiarization phase</i>	
Python, python, python, python, python, python, python, python, python, python	
<i>Test phase</i>	
1. Met een python moet je altijd voorzichtig zijn. (You should always be careful with a python.)	
2. Gelukkig vangt de lange hofnar hem nog op. (The tall court jester will look after him fortunately.)	
3. Zonder een hofnar lacht er nooit iemand hier. (Without a court jester no one here would ever laugh.)	
4. Dat is een lange python met scherpe tanden. (That is a long python with sharp teeth.)	
5. De hofnar maakt weer eens rare grappen. (The court jester sometimes makes weird jokes.)	
6. De koning hoort de boze hofnar vallen. (The king hears the nasty court jester falling.)	
7. De python ziet er nogal gevaarlijk uit. (The python looks rather dangerous.)	
8. Daar zie ik een boze python liggen. (I can see a nasty python lying there.)	

necessarily recognized by infants as undivided wholes or whether recognition of a previously heard word in continuous speech can be initiated on the basis of part of the word.

Little is known as yet about the ERP responses corresponding to the beginnings of word recognition in infants. The Mismatch Negativity (MMN) paradigm, a passive oddball paradigm in which an unexpected change in a series of stimuli usually results in a negative-going increase in ERP amplitude, has proven to be an extremely useful method for studying auditory discrimination of tones, phonemes or syllables [4], and studies have also been conducted on discrimination of (isolated) pseudo-words (e.g., in 4- and 5-month-old infants: see Ref. [17]) and (isolated) words (4- to 7-year-old children: see Ref. [11]). However, this type of paradigm is less optimal for answering the current research question, for which more complex stimuli, for example, spoken sentences, are required. To study word recognition from continuous speech, we need a paradigm in which it is possible to present (both isolated words and) full sentences.

For this, we exploited an ERP paradigm previously used in memory research [14], but in a novel way. The ERP procedure that we used had separate Familiarization and Test phases, on analogy with the two-phase HPP studies. In the Familiarization phase, we presented our participants, 28 prelinguistic 10-month-old Dutch infants, with lists of isolated Dutch words. Each list consisted of 10 tokens of the same two-syllable words (e.g., *python* ‘python’, *hofnar* ‘court jester’). The words were low in frequency and hence unlikely to be known by 10-month-olds. All had stress on the first syllable; this is a very common word structure in both English and Dutch [2,5], and the headturn preference response has been consistently observed for such words in

both languages [7,10,12]. The 10 tokens of any given word were each pronounced separately, so no two were identical, and each was spoken in the animated manner typical of speech directed to infants; the utterances depicted in Fig. 1 are taken from our materials. The Test phase, which immediately followed each word list, comprised eight sentences, four of which contained the familiarized words and four of which contained novel words (see Table 1 for an example of a Familiarization phase and a Test phase).

2. Materials and methods

2.1. Participants

Twenty-eight Dutch 10-month-old infants (mean age 308 days, range 288–320 days; 10 female) participated. Sixteen additional infants were tested but excluded from further analyses because they failed to complete enough of the experiment or because the data were too noisy due to movement artifact. The parent(s) gave informed consent for participation of their infant in the study. All infants came from monolingual Dutch families without left-handedness in the immediate family. No neurological or language problems were present in the immediate family. There were no serious complications during pregnancy. All infants were carried to full term except for one infant who was born 5 weeks pre-term. No hearing or sight problems were reported by the parents.

2.2. Materials

Forty low frequency, two-syllable nouns (from here onwards: target words; see Table 2 for a list of all target words) with a strong-weak stress pattern (that is, stress on the first syllable) were selected from the CELEX Dutch

Table 2
The 40 Dutch stimulus nouns

kiwi	sheriff
sitar	knolzwam
hommel	mammoet
monnik	sultan
zwaluw	viking
pelgrim	mosterd
maestro	parka
logo	kajak
tuba	medley
krokus	slede
serre	krekel
gondel	otter
orka	emoe
klamboe	toffee
sandwich	metro
drummer	hinde
ketjap	tabberd
pudding	sauna
hofnar	python
fakir	poema

lexical database [2]. Sets of four sentences containing each word were constructed. The target words and their component syllables were distinctive and unlikely to be familiar to the infants (e.g., *python* ‘python’, *hofnar* ‘court jester’). Position in the sentence and word preceding the target word were matched across sets. Words and sentences were recorded in a sound-attenuating booth onto digital audiotape by a native Dutch female speaker in animated child-directed speech, sampled at 16 kHz mono to disk and edited using a speech waveform editor. The mean duration of the target words was 710 ms (range: 363–1269 ms) in isolation and 721 ms (range: 224–1046 ms) in the sentences; mean sentence duration was 4082 ms (range: 2697–5839 ms). The onsets of the target words within the sentence contexts were labeled using a speech editing software package. Onsets were determined by a visual and auditory inspection of the speech signals.

2.3. Procedure

The experiment comprised 20 experimental blocks, each consisting of ten tokens of the same strong-weak word (familiarization stimuli) followed by eight randomized sentences; four of these sentences contained the familiarized word (Familiar condition), while the others contained a non-familiarized strong-weak word (Unfamiliar condition). Four versions of the experiment were compiled, counterbalancing familiarization token (i.e., each Familiarization list contained half of the target words) and order in which the experimental blocks were presented (i.e., one normal order and one reversed order). Every target word thus occurred in the Test phase for half of the infants as a Familiarized word and for the other half of the infants as an Unfamiliar control, and every infant heard both Familiar and Unfamiliar words. Each version of the experiment was presented to seven infants.

During the experiment, the infant sat in a child seat in a sound-attenuating booth. Approximately 1.5 m. in front of the child were three speakers, which presented the stimuli, and a computer screen continuously showing a colorful, moving, transforming object, which was not synchronized with the auditory stimuli. The child was allowed to play with a small silent toy during the experiment. The parent sat next to the child, listening to masking music through closed-ear headphones (Sennheiser HD 270). Since the experiment was too long for most infants, we presented as many of the 20 blocks as possible until the child became too distracted to continue. Each block took approximately 1.6 min, with approximately 2.5 s of silence between isolated words and 4.2 s between sentences. Breaks were taken when necessary. No subject heard fewer than nine blocks.

2.4. EEG recordings

EEG was recorded with an infant-size BrainCap with 27 Ag/AgCl sintered ring electrodes. Twenty-one electrodes

were placed according to the American Electroencephalographic Society 10% standard system (midline: Fz, FCz, Cz, Pz, Oz; frontal: F7, F8, F3, F4; fronto-temporal: FT7, FT8; fronto-central: FC3, FC4; central: C3, C4; centro-parietal: CP3, CP4; parietal: P3, P4; and occipital: PO7, PO8) [1]. Six electrodes were placed bilaterally on non-standard positions: a temporal pair (LT and RT) at 33% of the interaural distance lateral to Cz, a temporo-parietal pair (LTP and RTP) at 30% of the interaural distance lateral to Cz and 13% of the inion–nasion distance posterior to Cz, and a parietal pair (LP and RP) midway between LTP/RTP and PO7/PO8. All electrodes were referenced to the left mastoid online. The EEG electrodes were re-referenced offline to linked mastoids. Vertical eye movements and blinks were monitored via a supra- to sub-orbital bipolar montage (vEOG), and horizontal eye movements via a right-to-left canthal bipolar montage (hEOG).

EEG and EOG data were recorded with a BrainAmp AC EEG amplifier using a bandpass of 0.1–30 Hz and a sample rate of 200 Hz. Impedances were below 10 k Ω for all electrodes. Individual trials were aligned offline 200 ms before the acoustic onset of the target words. Four parietal and occipital electrodes (Pz, Oz, PO7, PO8) were excluded from analysis due to excessive artifact. EEG signal at the remaining 23 electrodes (three midline and 20 lateral electrodes) was screened for artifact from 200 ms before to 800 ms after acoustic onset of the critical word. Trials with artifacts were rejected (isolated words: 68%, words in sentences: 65%). This high percentage of artifact, mainly resulting from head movement, is normal in baby studies (for comparison, see Ref. [13]). For each subject average waveforms were calculated for each condition in this window. The grand average waveforms were calculated by averaging the subject average waveforms. The mean number of trials per condition per subject in the Familiarization phase was 8.3 for the unfamiliar words (i.e., word position 1/2; range 2–19) and 7.4 for the familiar words (i.e., word position 9/10; range 1–17). The total number of trials in the grand average was 231 for the familiar words and 207 for the unfamiliar words. In the Test phase, the mean number of trials per subject was 18.6 for the Unfamiliar condition (range: 12–34 trials) and 17.4 for the Familiar condition (range: 10–34 trials). The total number of trials in the grand average was 521 for the Unfamiliar conditions and 488 for the Familiar condition. Overall analyses were conducted over the subject averages across the 20 lateral electrodes, except where otherwise specified.

3. Results

We examined the ERP response during familiarization in order to establish criteria for the recognition response we could expect during the Test phase. Thus, we first analyzed the ERP response across the ten trials of the Familiarization phase. ERP responses were calculated for each two

successive trials (that is, word positions: e.g., position 1/2 is the average of the words in position 1 and 2). The grand average waveform (Fig. 2a) shows an extended positivity for position 1/2, starting at about 200 ms, mostly on frontal and fronto-central electrodes. Position 5/6 already shows a reduction of this positivity, but by position 9/10, there is an even further reduced positivity. Fig. 2b shows the mean amplitude in the window 200–500 ms for each two successive word positions. Positivity clearly diminishes with familiarization. The mean amplitudes for word positions 1/2 and 9/10, in the window 200–500 ms from word onset, were analyzed with repeated measures analysis of variance statistics (ANOVAs), with Familiarity (positions

1/2 versus 9/10) and Quadrant of the brain (right vs. left and frontal vs. posterior) as independent variables. All tests used the Huynh–Feldt epsilon correction. Familiar words were less positive than unfamiliar words (significant main effect of Familiarity: $F_{1,27} = 9.85, P = 0.004$). This Familiarity effect differed across quadrants (significant Familiarity x Quadrant interaction: $F_{1,27} = 6.34, P = 0.002$). In separate analyses by quadrant, the Familiarity effect was significant in the left and right frontal quadrants ($F_{1,27} = 19.45, P < 0.001$; $F_{1,27} = 10.84, P = 0.003$, respectively). Fig. 2c illustrates the distribution of this effect in an isovoltage plot. Next, we determined the exact onset of the Familiarity effect by examining the difference waveform (word position 9/10-

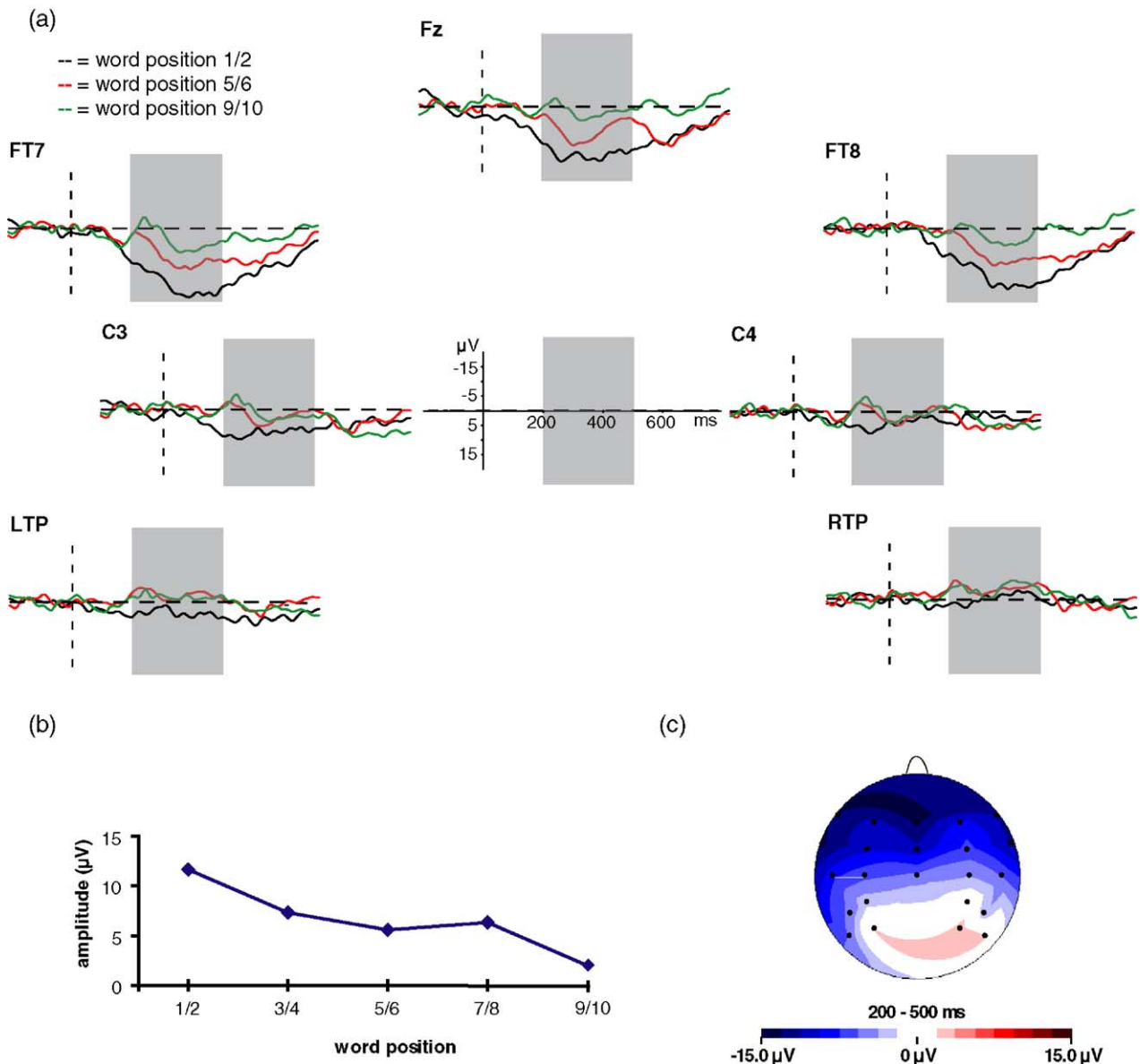


Fig. 2. Results of the Familiarization phase. (a) Event-related brain potentials (ERPs) to word positions 1/2, 5/6 and 9/10 at seven representative electrode positions Fz, FT7, FT8, C3, C4, LTP and RTP; negativity is plotted upwards. The grey area indicates the time window from 200 to 500 ms from word onset. (b) Mean amplitude (µV) per word position (i.e., 1/2, 3/4, 5/6, 7/8, 9/10) from 200 to 500 ms over the frontal, fronto-temporal and fronto-central electrodes. (c) Isovoltage plot of the familiarization effect in the Familiarization phase. The map is based on the difference waveform calculated for 23 electrodes by subtracting the ERP of word position 1/2 (unfamiliar words) from the ERP of word position 9/10 (familiar words) in the 200–500 ms latency range.

word position 1/2) and testing the difference from 0 (with two-tailed t tests) on consecutive 50 ms bins that shifted in steps of 10 ms (i.e., 0–50 ms, 10–60 ms, etc.) [3]. Significance ($P < 0.05$) on 5 consecutive bins was taken as evidence for onset of the Familiarity effect. This criterion was reached in the latency range of 160–190 ms for 16 (of 23) electrode sites in both hemispheres (14 frontal, fronto-central, central, fronto-temporal and temporal electrodes in both hemispheres and two parietal electrodes on the left hemisphere (F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, LT, C3, C4, RT, LTP, CP3; $P < 0.05$)). Thus, the Familiarization phase has produced a clear Familiarity effect—reduced positivity with increasing familiarity—Starting very

early on in the word (at about 160 ms) and spanning most of the word's duration.

We next examined the Test phase, in which infants had to segment words from continuous speech, in the light of this finding. Fig. 3a shows grand average waveforms of the target words in the sentences with words that had been presented in the Familiarization phase (Familiar words) contrasted with the same words when they had not been presented in the Familiarization phase (Unfamiliar words). Familiar words showed a greater negative deflection from 350 to 500 ms than did Unfamiliar words; this effect is in the same direction as in the Familiarization phase. This response was observed over the left hemisphere, but not

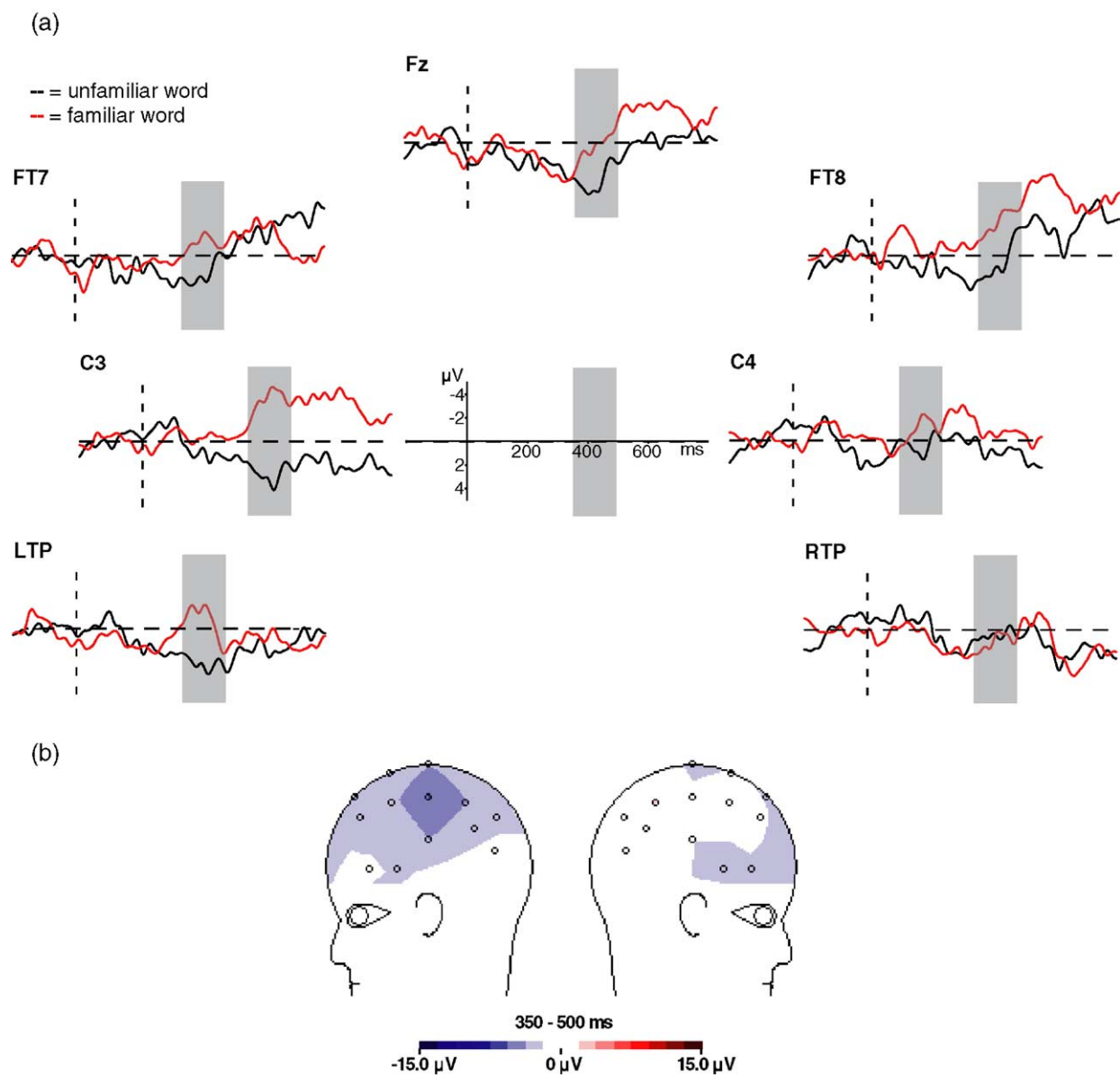


Fig. 3. Results of the Test phase. (a) Event-related brain potentials (ERPs) to the unfamiliar words and the familiar words in the sentences at seven representative electrode positions Fz, FT7, FT8, C3, C4, LTP and RTP; negativity is plotted upwards. The grey area indicates the time window from 350 to 500 ms from word onset. (b) Isovoltage plots of the familiarization effect in the Test phase. The map is based on the difference waveform calculated for 23 electrodes by subtracting the ERP of the unfamiliar words from the ERP of the familiar words in the 350–500 ms latency range.

over the right hemisphere (Fig. 3b). Mean amplitudes over the 350–500 ms time window were calculated and analyzed with repeated measures ANOVAs. Again, all tests used the Huynh–Feldt epsilon correction. We found no overall effect of Familiarity (20 lateral electrodes, $P > 0.05$), but we did find a significant interaction for Familiarity by Hemisphere ($F_{1,27} = 5.01$, $P = 0.034$). In separate analyses by hemisphere, we found a significant Familiarity effect over the left hemisphere ($F_{1,27} = 4.84$, $P = 0.037$), but not over the right ($P > 0.05$). We analyzed the onset of the Familiarity effect in the difference waveform (familiar words–unfamiliar words) in the same manner as for the Familiarization-phase responses. The criterion of $P < 0.05$ on five consecutive 50-ms bins was reached in the latency range of 340–370 ms for four left temporo-parietal electrode sites (C3, LT, LTP and CP3, $P < 0.05$). This response to the Test materials is, as in the Familiarization phase, positive-going, but begins later.

4. Discussion

The method we have developed has allowed us to see a cortical effect of word Familiarity in the 10-month-old's brain. The effect takes the form of a reduced positivity with increasing familiarity. In the Familiarization phase, we observed that the effect started very early on in the word (at about 160 ms). The two-syllable words were on average 710 ms long, so the Familiarity response started *while* the infants were hearing the early parts of the words. In the Test phase, we observed further evidence of a recognition response to the words that had been presented in the Familiarization phase. Here, the 10-month-olds heard every one of these words in a position internal to the sentences in running speech, and in no case was there a pause at the boundary of the critical word. Yet these infants, for whom the utterances were presumably meaningless, initiated the recognition response to the familiar words within half a second. Initiation of the response was about 180 ms later in the Test phase, in which the words occurred in the sentences, than in the Familiarization phase, in which the words occurred in isolation. Words in isolation are preceded by silence and their onsets are therefore abundantly clear. Words in a sentence are preceded by speech and determining the point of onset, as Fig. 1 demonstrated, is non-trivial. The listener must recognize, among the other speech sounds that are being processed, the familiar sounds which correspond to the known word. That extraction of this familiar sequence is not without cost is, then, represented by the 180 ms difference which we assume represents the cost of segmenting the words from the surrounding continuous speech.

Both ERP responses, in the Familiarization phase and in the Test phase, represent repetition responses; in the Familiarization phase, the response is immediate repetition of a token of the same isolated word, in the Test phase, it

reflects a comparison between repetition and no repetition in the context of a spoken sentence. Even though both effects are observed in the same direction, that is, a decrease in positivity with increased familiarity of the words, the different distributions of the effects (frontal, fronto-central, fronto-temporal in the Familiarization phase vs. left lateralized in the Test phase) suggests that partly different processes underlie them, and hence that (partly) different generators in the brain are responsible for these two effects. This is not surprising, since the Familiarization phase requires the infant merely to recognize different tokens of the same word, whereas the Test phase requires the infant to segment the word from continuous speech and recognize it as a familiar word. Recognizing the word in continuous speech is not the same as recognizing it in isolation. So the additional processing is visible in the difference both in latency and in the difference in distribution of the familiarization effects.

Previous studies by Jusczyk et al. [7–10,12] had shown with behavioral measures (the HPP paradigm) that infants in this age group prefer to listen to speech containing some words with which they had been familiarized in isolation over speech made up of only unfamiliar words. Our study relates this preference previously found in HPP studies to a precise and rapid cortical recognition response to those familiarized words embedded in continuous speech. The infant listeners achieve segmentation from the preceding context and launch the recognition response, all within the time-course of the word's delivery. The mean length of the two-syllable words in sentences was 721 ms, and yet the infants initiated the segmentation and recognition process by 340–370 ms. Thus, the process began by the end of approximately the first (stressed) syllable. In other words, infants cannot be matching whole-word templates against the input. They must be accessing memory representations that have sufficient internal structure for the initial portion of these words in the speech context to be matched to the initial portion in the representation constructed during familiarization.

In our experiment, we may assume that no semantic representation was activated in the 10-month-olds' memory when a newly familiar word form was re-encountered. Thus, ERP studies with adults or older children hearing words and nonwords [6,13,15] provide no guide for the present case because word recognition always involves activation of meaning when listeners already possess a vocabulary. Since, in our study, the acoustic tokens representing the word forms varied substantially, and sound very different in continuous speech than they do in isolation, the 10-month-olds have apparently acquired the capacity to generalize across different acoustic tokens and to categorize them at a more abstract phonological level. Exactly which cues in the continuous speech signal are used by the infants to trigger the segmentation and word recognition process is still an open issue and a topic for further research. One likely candidate worthy of further exploration might, for instance, be the stress pattern. In a stress-based language, the syllables that

carry stress might be units that the infant uses to start up segmentation and recognition; this suggestion has been made for English [10], and Dutch, like English, is stress-based. In any case, we suggest that our paradigm takes the study of cortical responses to speech in infancy a step further, in that it is now possible to investigate the previously intractable issue of infants' brain responses to word recognition in continuous speech.

For infants to construct an initial vocabulary and begin to speak, they must first be able to recognize word forms on repeated occurrence despite the inevitable variability in the pronunciations. And because much of the speech they hear consists of continuous multi-word utterances [16], they must develop the ability to extract individual word forms from continuous speech. Whether infants segment the whole word or only the salient first part is a topic for further research. However, our results clearly show that this ability is already so finely tuned by 10 months of age that infants can start segmenting and recognizing the onset of a familiar word embedded in continuous speech within half a second.

Acknowledgments

This research was supported by a SPINOZA grant from the Netherlands Organization for Scientific Research (NWO) to AC. We thank Dan Swingley and Marlies Wassenaar for useful assistance, Herb Clark, James McQueen, Elizabeth Johnson and two anonymous reviewers for comments on the manuscript, and Dennis Pasveer for making Fig. 1.

References

- [1] American Electroencephalographic Society, Guidelines for standard electrode position nomenclature, *J. Clin. Neurophysiol.* 11 (1994) 111–113.
- [2] R.H. Baayen, R. Piepenbrock, H. Van Rijn, The CELEX lexical database [CD-ROM], Linguistic Data Consortium, University of Pennsylvania, Philadelphia, 1993.
- [3] J.A. van Berkum, C.M. Brown, P. Hagoort, Early referential context effects in sentence processing: evidence from event-related brain potentials, *J. Mem. Lang.* 41 (1999) 147–182.
- [4] M. Cheour, P.H.T. Leppänen, N. Kraus, Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children, *Clin. Neurophysiol.* 111 (2000) 4–16.
- [5] A. Cutler, D.M. Carter, The predominance of strong initial syllables in the English vocabulary, *Comput. Speech Lang.* 2 (1987) 133–142.
- [6] H.J. Heinze, T.F. Münte, G.R. Mangun (Eds.), *Cognitive Electrophysiology*, Birkhäuser, Boston, 1994.
- [7] D.M. Houston, P.W. Jusczyk, C. Kuijpers, R. Coolen, A. Cutler, Cross-language word segmentation by 9-month-olds, *Psychon. Bull. Rev.* 3 (2000) 504–509.
- [8] P.W. Jusczyk, How infants begin to extract words from speech, *Trends Cogn. Sci.* 3 (1999) 323–328.
- [9] P.W. Jusczyk, R.N. Aslin, Infants' detection of the sound patterns of words in fluent speech, *Cogn. Psychol.* 29 (1995) 1–23.
- [10] P.W. Jusczyk, D.M. Houston, M. Newsome, The beginnings of word segmentation in English-learning infants, *Cogn. Psychol.* 39 (1999) 159–207.
- [11] P. Korpilahti, C.M. Krause, I. Holopainen, A. Heikki Lang, Early and late mismatch negativity elicited by words and speech-like stimuli in children, *Brain Lang.* 76 (2001) 332–339.
- [12] C. Kuijpers, R. Coolen, D. Houston, A. Cutler, Using the head-Turning technique to explore cross-linguistic performance differences, in: C. Rovee-Collier, L. Lipsitt, H. Hayne (Eds.), *Adv. Infancy Res.*, vol. 12, Ablex, London, 1998, pp. 205–220.
- [13] D.L. Mills, S.A. Coffey-Corina, H.J. Neville, Language acquisition and cerebral specialization in 20-month-old infants, *J. Cogn. Neurosci.* 5 (1993) 317–334.
- [14] M. Rugg, D. Doyle, Event-related potentials and stimulus repetition in direct and indirect tests of memory, in: H.J. Heinze, T.F. Münte, G.R. Mangun (Eds.), *Cognitive Electrophysiology*, Birkhäuser, Boston, 1994.
- [15] M.D. Rugg, M.C. Doyle, J.S. Holdstock, Modulation of event-related brain potentials by word repetition: effects of local context, *Psychophysiology* 31 (1994) 447–459.
- [16] J. Van de Weijer, Language input for word discovery [doctoral dissertation], MPI Ser. Psycholinguist., vol. 9, MPI, Nijmegen, 1999.
- [17] C. Weber, A. Hahne, M. Friedrich, A.D. Friederici, Discrimination of word stress in early infant perception: electrophysiological evidence, *Cogn. Brain Res.* 18 (2003) 149–161.