Continuous-speech segmentation at the beginning of language acquisition: electrophysiological evidence

ISBN: 978-90-76203-29-4

Design MPI Series in Psycholinguistics: Linda van den Akker & Inge Doering
Cover illustration: 'Ísland' by Valesca Kooijman, edited by Frank Kooij
Printed & bound by Ponsen & Looijen b.v., Wageningen

Continuous-speech segmentation at the beginning of language acquisition: electrophysiological evidence

Een wetenschappelijke proeve op het gebied van de

Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op het gezag van rector magnificus **prof. mr. S.C.J.J. Kortmann**
volgens besluit van het College van Decanen
in het openbaar te verdedigen op **maandag 22 oktober 2007**
om 13:30 uur precies

door

Valesca Madalla Kooijman
geboren op **26 januari 1974**
te **Ede**

Promotoren:
Prof. dr. Anne Cutler
Prof. dr. Peter Hagoort

Manuscriptcommissie:
Prof. dr. Harold Bekkering
Dr. Miranda van Turennout
Prof. dr. Richard N. Aslin (University of Rochester, Rochester, USA)

The research reported in this thesis was supported by the Spinoza grant 'Native and Non-native Listening' from the Netherlands Organization for Scientific Research (NWO) to A. Cutler and by NWO grant 051.04.031 to P. Hagoort

Dankwoord

Zes jaar na het begin van mijn promotieproject is het dan eindelijk zover. Tijd om het dankwoord te schrijven! Zonder de hulp van vele anderen was dit proefschrift niet tot stand gekomen.

Allereerst wil ik mijn promotor Anne Cutler en copromotor Peter Hagoort bedanken. Anne wil ik met name bedanken voor haar waardevolle lessen over wetenschappelijk schrijven in het Engels, en Peter voor mijn gedegen opleiding tot EEG onderzoeker.

Verder wil ik Dan Swingley, Elizabeth Johnson, Herb Clark, Dick Aslin, Marlies Wassenaar, Angela Khadar, John Nagengast, Johan Weustink, Erik van den Boogert en mijn stagiaires Floor, Tineke, Carolien, en Caroline bedanken voor hun bijdrage aan mijn proefschrift. En natuurlijk alle ouders en kindjes, en andere deelnemers aan mijn onderzoek. Bedankt voor jullie tijd en inzet!

Uiteraard hebben de vele uurtjes theeleuten op het werk geholpen dit project vol te houden. Marieke, Marlies, Oliver, Sandra V., Cathelijne en Sandra H., en vele anderen, ontzettend bedankt voor alle gezelligheid en het van tijd tot tijd broodnodige luisterend oor.

Op het werk, maar vooral ook daarbuiten, hebben mijn vrienden me steeds bijgestaan. Dannie, Annet, Heleen, Carolien, Marjolijn, Heidi en Eric, heel erg bedankt voor jullie steun in stressvolle tijden maar natuurlijk ook voor al het plezier en de gezelligheid de afgelopen jaren. Rudy, Rieky, Rob, Kat, Jari, Jeroen en Marieke, Evelien, en Albert, bedankt voor jullie belangstelling en vriendschap.

Bovenal wil ik mijn ouders Jan en Stella, mijn zussen (en paranimfen!) Tabitha en Karucha, en Oema bedanken. Zonder jullie was ik nooit zo ver gekomen. Bedankt voor alles!

Tot slot, lieve Frank, dank je wel voor al je steun en liefde, en alle boten en brugjes. Zonder jou waren de afgelopen anderhalf jaar niet half zo fijn geweest.

Contents

CONTENTS

Introduction

Chapter 1

This chapter is based in part on the book chapter Kooijman, V., Johnson, E.K., & Cutler, A., in press. Reflections on reflections of infant word recognition. In: A. Friederici and G. Thierry (Eds.), Trends in Language Acquisition Research, Amsterdam: John Benjamins Publishing.

Acquiring a language is a great accomplishment, not only for adults, but maybe even more so for infants. Starting even before birth, infants get acquainted with their native language by hearing their mother's voice. From here onwards, language development proceeds extremely rapidly, leading to a more or less set language system at the age of three. This thesis deals with one aspect of this amazing accomplishment, known as *word segmentation*. Word segmentation refers to the ability to divide the speech stream into its component words. Infants acquire this ability in the second half of the first year of life. In the experimental chapters of this thesis, several Event Related Brain Potential (ERP) studies on word segmentation will be discussed. In this introductory chapter, a brief overview of early language development and the different research methodologies is given, as well as a summary of brain development. Subsequently, the word segmentation problem in both adults and infants is described, as well as the Headturn Preference Procedure (HPP), a behavioral method particularly suitable to study behavior in infants. Next, ERP and other neuroimaging techniques are described. Finally, an outline of the remainder of this thesis is given.

LANGUAGE DEVELOPMENT EARLY IN LIFE

Native language acquisition can be roughly divided into three stages. In the first year of life, infants learn the sound structure of the native language. Sensitivity to the native language phonology increases at a rapid pace, whereas sensitivity to non-native phonology reduces. Early in the second year of life, comprehension and production of the language become increasingly important. Later in the second year, and continuing into the third year of life, the vocabulary spurt and a vast increase in knowledge of syntactic structure play a major role. By the end of the third year, the native language is more or less stable and in place. From here onwards, language development mostly consists of increasing fluency in the use of the native language. (For detailed overviews on different aspects of language acquisition, see Bates, Thal, Finlay, & Clancy, 2002; Clark, 2004; Kuhl, 2004; Peperkamp, 2003; Werker, 2003; Werker & Tees, 1999.)

The first year of life

Although ERP studies with infants are becoming increasingly popular, the bulk of what we know about language acquisition comes from behavioral studies. The commonest behavioral testing methodologies have used the rate or duration of simple behavioral responses, such as sucking on a pacifier or looking at a visual stimulus associated with an auditory signal, as the indirect measures of developing speech perception and processing abilities. Creative use of these testing methodologies has uncovered remarkably sophisticated speech perception skills in preverbal infants. The High Amplitude Sucking Paradigm, for example, which uses sucking rate as a dependent measure of speech preferences and discriminatory abilities, works well with infants up to two months of age (Jusczyk, 1985; Sameroff, 1967). Research using this paradigm has demonstrated that infants begin laying a foundation for language acquisition even before birth. Newborns prefer to listen to their mother’s native tongue over other languages (e.g. English-learning infants prefer to listen to English over Spanish; Moon, Cooper, & Fifer, 1993). They also show recognition of voices (DeCasper & Fifer, 1980) and of stories heard before birth (DeCasper & Spence, 1986), and they discriminate phoneme contrasts (Eimas, Siqueland, Jusczyk, & Vigorito, 1971).

Of course, newborns are still far removed from linguistic competence. Their phoneme discrimination skills reflect their auditory abilities, not their use of linguistic experience; they can as well discriminate phonetic contrasts which do not appear in the maternal language as those that do (Aslin, Jusczyk, & Pisoni, 1998; Werker & Tees, 1984; 1999). At two months of age, likewise, English-learning infants cannot yet perceive the difference between their own language and the rhythmically similar Dutch (Christophe & Morton, 1998). However, speech processing skills develop rapidly during the first year of life, as research using other procedures more suited to testing older infants, such as the Conditioned Headturn Procedure (CHP) and the Headturn Preference Procedure (HPP), has demonstrated. These procedures make use of the infants' natural inclination to turn their heads in the direction of the sounds they hear. The infants' head turn in the direction of auditory stimuli is then interpreted as listening time. The longer listening time to one type of stimulus over another indicates a preference (Fernald, 1985; also see Werker, Polka, & Pegg, 1997, and the section '*The Headturn Preference Procedure and early word segmentation*' of this chapter). Such paradigms have been used to show that by four months, infants recognize their own name (Mandel, Jusczyk, & Pisoni, 1995) and discriminate between their native language and other rhythmically similar languages (Bosch & Sebastián-Gallés, 1997). By five months, infants are so familiar with the prosodic structure of their native language that they can even discriminate between two dialects of their native language – thus American infants discriminate between American and British English (Nazzi, Jusczyk, & Johnson, 2000). Sensitivity to language-specific vowel patterns emerges by six months of age (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992), and language-specific consonant perception is well in place before infants reach their first birthday (Werker et al., 1984; 1999). First evidence of rudimentary word segmentation and comprehension skills has been observed between six and seven and a half months of age (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Jusczyk & Aslin, 1995; Tincoff & Jusczyk, 1999). The ability to coordinate more than one source of information arises between eight and ten months of age (Jusczyk, 1999; Morgan & Saffran, 1995; Werker et al., 1999). In speech perception at about nine months of age, infants are able to coordinate several phonetic cues. This is very important for, among other things, word segmentation. This 'sudden' ability to deal with more than one source of information is also seen in other areas of development, such as attention and memory. Although some researchers claim a language-specific account of the change in language perception skills, a more general underlying change in the use of information may be a more likely explanation (Werker et al., 1999). Word segmentation and comprehension skills continue to develop at an impressive rate in the last months of the first year (Hollich, Hirsch-Pasek, & Golinkoff, 2000; Jusczyk, Houston, & Newsome, 1999). In addition, babbling becomes more language-specific and infants start producing their first words (Bates et. al., 2002; Werker et al., 1999).

The second year of life

In the second year of life, word comprehension and production as well as grammatical learning increase rapidly. Roughly between 11 and 13 months of age, infants learn to comprehend about 50 words of their native language. First word production and object naming is initiated early in the second year. General cognitive skills such as joint reference and attention play an important role in learning these language-specific skills (Bates et al., 2002, Werker et al., 1999).

Word-picture matching tasks are commonly used in the second year of life, for example, to study the representation of phonetic detail in the initial lexicon. Swingley and Aslin (2000) used such a word-picture matching task with familiar words to show a highly detailed representation of some words in 14-month-olds. Infants looked considerably longer at a picture of a baby while listening to <*baby*> than when listening to the very similar non-word <*vaby*>. At

17 months of age, but not at 14 months, infants can discriminate between phonetically highly similar words associated with a new object (Werker, Fennell, Corcoran, & Stager, 2002). Thus, some level of phonetic detail seems to be present in the early lexicon and it increases rapidly over the next few months. At about 18 months of age, supposedly, the vocabulary spurt takes place, characterized by a sudden and fast increase in word production (e.g., Bates et al., 2002). However, this sudden spurt has recently been debated and a general increase in word learning and production throughout childhood has been suggested instead (Bloom, 2000; Ganger & Brent, 2004). At about 20 months of age the production of word combinations begins. From here onwards, a fast increase in grammatical learning can be seen as well as the production of longer word combinations (Bates et al., 2002; Werker et al., 1999), and infants are by this time well on their way to adult-like language comprehension and production.

Of course, in addition to cognitive development, the infant's brain develops at a rapid pace as well. Relatively little, however, is known about early brain development. The next section gives a brief summary.

Brain development

Before birth, the fetus' brain shows an impressive level of growth. All cells are generated by the third trimester after gestation and the major nerve pathways are in place. There is even some level of learning possible in the last weeks before birth (Bates et al., 2002). Nevertheless, brain development does not reach its mature level until late into the second decade of life. For example, myelination, i.e., the increase in fatty sheath that surrounds the neuronal pathways, continues until years after birth (Pujol et al., 2006; Uylings, 2006). The myelin sheath helps increase the speed and efficiency of signal transmission through the axons. Dendritic growth (Mrzljak, Uylings, Van Eden, & Judas, 1990; Uylings, 2006) and the development of cortical folding (i.e., the development of tertiary convolutions; Toro & Burnod, 2005) continue well into the first year after birth. Synaptogenesis as well as specialization of brain areas through a decline in the number of cells (i.e., apoptosis), also both still take place after birth (Bates et al, 2002). These processes seem to take place in parallel and in waves in different areas of the brain. A longitudinal MRI study with older children (age 4 to 22) showed a linear increase in cortical white matter, and a non-linear increase and decrease in cortical grey matter varying per area of the brain (Gield et al., 1999). Frontal and parietal grey matter reached its peak volume around age 10-12, and showed a decrease during puberty. Temporal grey matter did not reach its peak volume until 16 years age with a decline afterwards, whereas only an increase but no decrease was seen in occipital areas. It is not clear yet which processes (e.g., changes in neuronal size, axonal or dendritic arborization) are involved in these changes in grey matter. A Positron Emision Tomography (PET) study on metabolic changes in the brain showed the highest metabolic rate in sensorimotor cortex, thalamus, brain stem and cerebellar vermis in infants younger than five weeks of age (Chugani, Phelps, & Mazziotta, 1987). At three months of age, metabolic rate had increased in parietal, temporal, and occipital areas. Frontal areas showed an increase in metabolism around six to eight months of age. Around two to four years, metabolic rate reached adult-like levels, but showed a decrease again around nine years of age, before it returned to adult levels at the end of adolescence.

Thus, brain development is not a linear process and continues well into the second decade of life. Its functional relationship with developing cognitive abilities is as yet not completely clear. In particular, the functional relationship between brain development and language skills needs further research. In the final sections of this chapter, neuroimaging techniques that can be used to study this relationship will be discussed. First however, the word segmentation problem will be described, as will the HPP, a behavioral method commonly used to study word segmentation.

WORD SEGMENTATION

Word segmentation in adults and infants

Hearing speech as a string of discrete words seems so effortless to adults listening to their native language that it is tempting to suspect that the speech signal unambiguously informs us where one word ends and the next begins. However, listening to an unfamiliar language or examining a spectrogram easily dispels this illusion. When we listen to an unfamiliar language, words seem to run together in a very fast manner; it is only in our own language that segmenting streams of speech into their component words is so easy (see also chapter 5 of this thesis). But in fact words run together in any language (Nazzi, Iakimova, Bertoncini, & de Schonen, in press). Figure 1 illustrates this with a Dutch eight-word sentence: *Die oude mosterd smaakt echt niet meer goed* 'that old mustard really doesn't taste good any more'. There are several silent portions in the speech stream, but even where these happen to occur between words, they have not arisen from pauses between the words: each such point just represents the closure of the speaker's mouth as a stop consonant (/d/, /t/, /k/, or the glottal stop separating successive vowels) has been uttered. The eight words are not demarcated by recurring word-boundary signals of any kind. This utterance was in fact spoken slowly and carefully in an infant-directed manner; most utterances in our everyday experience proceed even faster and weld the separate words even more closely together than we see here.

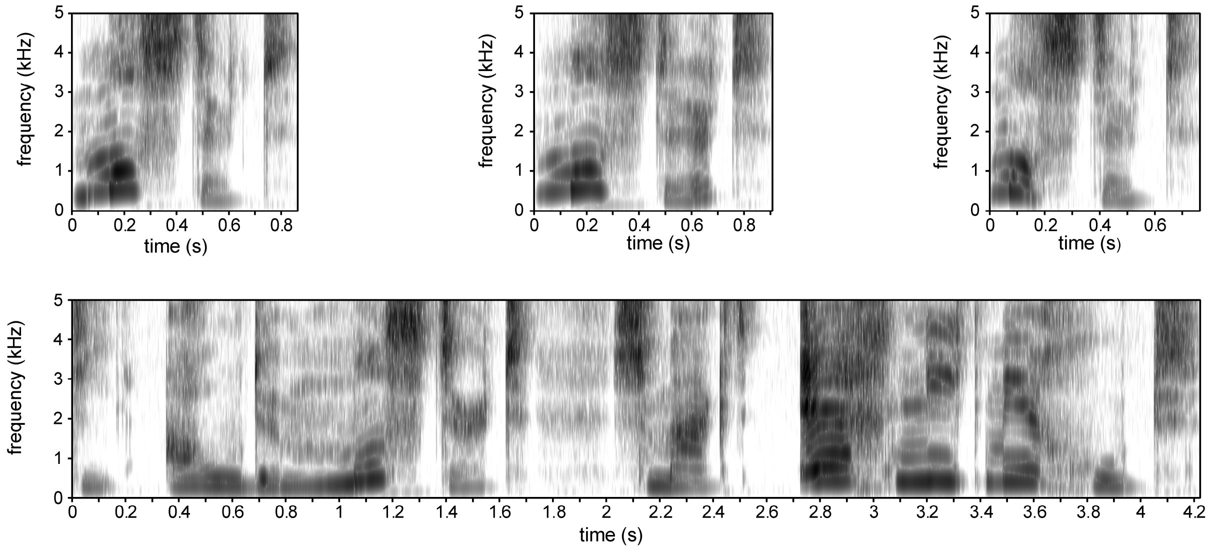

Figure 1: Spectrograms of Dutch words and a sentence. *Above, three spectrograms of the Dutch word 'mosterd' (mustard), produced in isolation in an infant-directed manner; below, a sentence 'Die oude mosterd smaakt echt niet meer goed' (That old mustard really doesn't taste good any more), produced in the same manner. 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 that the three word tokens differ in duration, from about 750 ms to about 900 ms, and also differ in spectral quality. The word mosterd in the sentence begins at about 0.78 on the time line and finishes at about 1.75.*

Why is it so easy to hear words in our native language? As it turns out, there are a myriad of cues to word boundaries which listeners can call upon, but these cues are probabilistic rather than being fully reliable; further, and most importantly, they are language-specific. Adults therefore exploit multiple cues to identify word boundaries in fluent speech, and the cues they use are determined by their native language experience (Cutler, 2001). Phonetic (i.e., properties of speech sounds) and phonotactic (i.e., possible phoneme combinations) regularities, the metrical stress pattern of the languages, and lexical information (i.e., information stored in the mental lexicon on candidate words and their grammatical and phonological properties) may all help the adult listening to their native language.

The role of these cues has been described in different models of word segmentation. In the earliest models of spoken-word recognition (Cole & Jakimik, 1978; Marlsen-Wilson & Welsh, 1978), words were simply processed sequentially. Segmentation occurred whenever enough of a word had been heard that its end could be identified; at that point, the next word would begin. However, competition accounts (McClelland & Elman, 1986; Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997; Gaskell & Marslen-Wilson, 1997) allow word segmentation to arise as a by-product of multiple simultaneous candidate words. Norris (1994), for example, proposed that spoken language activates lexical candidates that are partly or fully congruent with the input. As the input proceeds, some candidates will continue to receive further activation, whereas others become more incongruent and their activation level reduces. The more activation a candidate word has, the more it is able to inhibit rival candidates. This competition between lexical candidates leads to victory for, and recognition of the correct words in the input and thus, indirectly, to segmentation.

Several other models, however, consider pre-lexical regularities in the language as the initial cues to finding word boundaries (Elman, 1990; Brent & Cartwright, 1996; Cairns, Shillcock, Chater, & Levy, 1997; Christiansen, Allen, & Seidenberg, 1998; Wolff, 1977). Brent and Cartwright (1996), for example, proposed a model in which the main cues to speech segmentation are phonotactic regularities. Such a model presupposes knowledge of distributional regularities of the native language. Several studies have shown that both adults (Cairns, et al., 1997; McQueen, 1998) and infants at eight months of age (Saffran, 2001) are sensitive to these regularities. It is therefore safe to assume this level of knowledge in both adults and older infants.

Word stress is another useful cue for word segmentation, at least in stress-based languages such as Dutch and English (Cutler & Norris, 1988). Since the majority of English content words begin with a stressed syllable (Cutler &

Carter, 1987), English listeners are biased to perceive stressed syllables as word onsets (Cutler & Butterfield, 1992). English listeners who tried to apply this strategy to segmentation of spoken French, Polish or Japanese, however, would have little luck extracting words from the speech stream.

Some segmentation models propose a combined role for both competition and pre-lexical cues. In a word-spotting experiment, Norris, McQueen and Cutler (1995) showed that the effect of prosody on word segmentation increases with the number of lexical candidates. This suggests that segmentation depends on more than one process. Mattys, White and Melhorn (2005) pitted different types of word segmentation cues against each other in a series of experiments and suggested a hierarchical model for adult word segmentation with lexical cues at the top, followed by statistical regularities and metrical stress. However, in a noisy environment, the hierarchy reverses and metrical stress is the most important cue.

Thus, although there is no consensus yet on how exactly adults extract word boundaries from speech, it is clear that segmenting words from speech is a trivial task for adults as they are able to combine many sources of information, and have years of experience listening to their native language. Learning how to find words for the first time, however, presents a much bigger challenge. Infants do not have a lexicon yet in their first year, and are therefore not able to use lexical information. Nevertheless, learning to segment words from speech in the first year of life is very important as is clear from Newman, Bernstein Ratner, Jusczyk, Jusczyk, and Dow's (2006) demonstration that relative ability to recognize discrete words in continuous speech before age one is directly predictive of vocabulary size at age two. It has been proposed that infants might solve the word segmentation problem by first learning words in isolation, and then subsequently recognizing these words in fluent speech (Bloomfield, 1933; Brent, 1999). However, the speech which infants hear in the first year of life consists predominantly of multiword utterances (Morgan, 1996; Van de Weijer, 1998; Woodward & Aslin, 1990), so it seems unlikely that hearing words in isolation could constitute the full explanation for how language learners first begin segmenting words from speech. It seems more likely that the onset of word segmentation is fueled by developing knowledge about the typical sound pattern of words, i.e., by exploitation of language-specific probabilistic cues like typical phonotactic patterns and word stress patterns (Mattys, Jusczyk, Luce, & Morgan, 1999; Saffran, 2001). Mattys et al. (1999) used a Headturn Preference Procedure (HPP) design to show that both phonotactic patterns and word stress patterns are important for word segmentation in nine-month-olds. However, pitting phonotactic sequences and prosodic cues (i.e., word stress) against each other in a word segmentation task showed a stronger role for prosody (also see Johnson & Jusczyk, 2001). Chapters 2, 3, and 4 of this thesis further address the role of metrical stress on word segmentation, using ERP measures. However, the HPP also has proven to be a very important tool for the study of early word segmentation and will be discussed in detail in the next section.

The Headturn Preference Procedure and early word segmentation

The development of the Headturn Preference Procedure (HPP; see Figure 2 for an illustration of the setup of the HPP) brought about great advances in understanding of when infants begin segmenting words from speech. Before the HPP was in widespread use, evidence from language production led researchers to conclude that four-year-olds still had not completely solved the word segmentation problem (Chaney, 1989; Chaney & Estin, 1987; Holden & MacGinitie, 1972; Huttenlocher, 1964; Tunmer, Bowey, & Grieve, 1983). At the same time, however, most studies of early syntactic development assumed that two- and three-year-olds were perceiving speech as a string of discrete words. In retrospect, this assumption does not seem unwarranted, especially since it seems only logical that children would need to learn to segment words from speech before they could build a large enough vocabulary to communicate their thoughts verbally. In other words, research on infant word segmentation lagged behind research on, for instance, phoneme and language discrimination.

One reason for the relative lag is that studying word segmentation presents methodological challenges. First, long stretches of speech must be presented. Second, there must be a measure of recognition rather than simply of discrimination or preference. The earliest widely used infant testing methodologies, such as the High Amplitude Sucking Procedure and the Visual Fixation Procedure, were unsuited to the study of word segmentation because they offered no recognition measure.

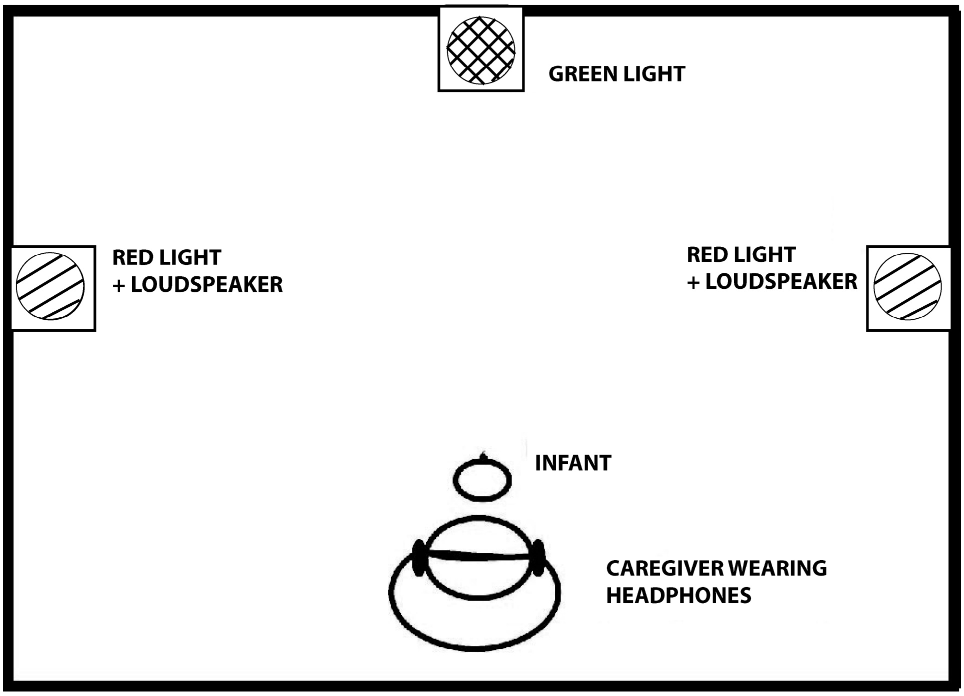

Figure 2: A HPP setup. *In a HPP experiment, the infant is seated on the caregiver's lap in a three-sided test booth. A green light is mounted in front of the child at eye level, and red lights are mounted on each side. During the experimental trials, the side lights are used to draw the infant's attention. If the infant makes a head turn in the right direction, stimuli start playing from a loudspeaker. The time spent looking in the direction of a stimulus type is interpreted as listening time. A longer listening time to one stimulus type over another is considered a preference.*

The first use of HPP was in a test of four-month-olds' preferences concerning adult- versus infant-directed speech (Fernald, 1985). In Fernald et al.'s experiment, infants sat facing forward on a parent's lap in the middle of a three-sided booth. A light was mounted at eye level in the center of each of the three walls of the booth. Speakers were hidden behind the lights on the two side walls; infant-directed speech (IDS) was played from one speaker and adult-directed speech (ADS) from the other. The green light on the front panel blinked at the onset of each trial. Once infants oriented towards the green light, it would immediately stop blinking and both of the side lights would begin blinking. Depending on which light the infants turned towards, they would hear either IDS or ADS. Headturns were observed by an experimenter out of view of the infant. Fernald et al. found that infants turned to the side from which IDS was played more often than they turned to the side from which ADS was played. Accordingly, they inferred that four-month-olds preferred to listen to IDS over ADS.

In this version of the HPP, the dependent measure was how often infants turned to the left versus right. In the first HPP study of word segmentation (Myers et al., 1996), the procedure was modified so that all stimulus types were played equally often from the left and right speaker, and the dependent measure was length of orientation time to speech from one side versus the other. The contrast in this study was between passages containing pauses inserted within words versus pauses inserted between words. Eleven-month-olds listened longer to the latter type of speech. Based on the assumption that infants prefer to listen to natural- over unnatural-sounding speech samples (see Jusczyk, 1997, for review), this study suggested that 11-month-olds have some concept of where word boundaries belong in speech. But this is not the best test of word segmentation abilities, since it is possible that the infants had simply noticed the unnatural disturbance of the pitch contour.

A better test of infants' word segmentation skills was devised by Jusczyk and Aslin (1995), who further modified HPP by adding a familiarization phase prior to the test phase (see also Kemler Nelson, et al., 1995). During the familiarization phase of Jusczyk and Aslin's study, 7.5-month-olds listened for 30 seconds to isolated repetitions of each of two words: *dog* and *cup* or *bike* and *feet*. In the test phase immediately following this familiarization, infants' length of orientation to test passages containing *dog*, *cup*, *bike*, and *feet* was measured. Infants familiarized with *bike* and *feet* listened longer to test passages containing *bike* and *feet*, while infants familiarized with *cup* and *dog* listened longer to passages with *cup* and *dog*. Six-month-olds tested with the same procedure and stimuli failed to demonstrate any listening preferences.

Jusczyk and Aslin accordingly concluded that infants begin segmenting words from speech some time between six and 7.5 months of age. Numerous subsequent segmentation studies with the two-part version of HPP have supported this finding (see Jusczyk, 1999, and Nazzi et al., in press, for reviews). In combination, these HPP studies have provided clear evidence that production studies underestimate the rate of development of infants' word segmentation ability. Production studies were inadequate to study early word segmentation for several reasons. First, they required a verbal response, which limited researchers to testing children who could already speak. Second, the tasks used to test children's ability to hear word boundaries were often quite complicated (e.g. repeating the words in an utterance in reverse order). The difficulty of these tasks is very likely to have masked younger children's ability to segment words from speech. Word segmentation abilities develop in the course of initial vocabulary building, and studies with the HPP made that clear.

Advantages and disadvantages of behavioral word segmentation measures

The HPP has many strengths as a testing methodology for research on word segmentation. First, it allows long stretches of speech to be presented in either the familiarization or test phase of the experiment; this is obviously an essential prerequisite for studying fluent speech processing. Indeed, recent studies have shown that HPP also works well with fluent speech in *both* familiarization and test phases (Seidl & Johnson, forthcoming; Soderstrom, Kemler Nelson, & Juscyk, 2005). Second, the dropout rate in HPP is relatively low compared to other testing methodologies. Third, HPP yields less variable data than some other methods, since looking-time measures are often based on 12 to 16 trials, rather than the two or four test trials commonly used, for example, in the Visual Fixation Procedure (however, see Houston & Horn, submitted, for discussion of an adapted version of the Visual Fixation Procedure allowing multiple test trials and providing results which are arguably suitable for individual subject analysis). Fourth, HPP is widely applicable; although it may be best suited for testing children between six and nine months of age, it has been shown to work well with children as young as four months or as old as 24 months. This is certainly useful, considering the protracted development of word segmentation abilities (e.g., see Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005). Fifth and finally, HPP does not require that infants be trained to focus on any particular aspect of the speech signal. Rather, in contrast to procedures like the Conditioned Headturn Procedure (CHT), it provides a measure of what infants naturally extract from the speech signal.

Like all infant testing methodologies, HPP has a few disadvantages too. As with other methods, it is hard to say whether performance in the laboratory is accurately representative of performance in the real world, where visual and auditory distractions are plentiful (see however, Newman, 2005). HPP is ill-suited to the study of individual variation, because a typical HPP study requires multiple subjects. Infants can become bored with the HPP procedure, and re-testing a child with the same procedure is not advisable. Finally, with particular importance for the case of word segmentation, HPP looking times do not reflect the temporal nature of the processing involved and requires a behavioral response. This may result in an underestimation of the cognitive competence of infants. Although the behavioral response is an expression of the level of processing an infant has reached, infants may be able to process certain types of information without being able to initiate a corresponding motor response.

Corresponding motor areas or connections to motor areas in the brain may not yet have matured significantly.

In adult word segmentation research, the temporal course of word processing has played an important role in understanding how words are recognized. Reaction time studies have revealed that many word candidates are simultaneously activated, and then compete for recognition (Norris, et al., 1995); the competition process is further modulated by explicit segmentation procedures which can be language-specific (e.g., attention to rhythmic structure; Cutler & Butterfield, 1992) or universal (e.g., rejection of activated words which would leave isolated consonants unaccounted for in the signal; Cutler, Demuth, & McQueen, 2002; Norris, et al., 1997). But the HPP effectively only tells us whether word segmentation has occurred, not how rapidly it has occurred. Evidence for the temporary activation of spurious word candidates, or information about the precise timing of online segmentation, cannot be found with HPP. Thus although we know that twelve-month-olds also fail to segment word candidates which would leave isolated consonants unaccounted for (Johnson, Jusczyk, Cutler, & Norris, 2003), the results of this study – summarized in Figure 3 – tell us only that segmentation has occurred in one condition and not in the other; they tell us nothing about the relative speed of word recognition which was addressed in the adult studies, let alone about the relative segmentation success for individual words in the passages or the performance of individual listeners.

It would certainly be advantageous if the fine-grained temporal course of word segmentation could also be studied in younger infants, who are just beginning to use their newly acquired knowledge about the sound structure of their native language to extract word forms from speech. Two procedures which appear more temporally sensitive than HPP each have limitations. First, eye-tracking procedures (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Swingley, Pinto, & Fernald, 1999) certainly offer a window onto the temporal course of children's processing; however, these procedures can only be used with children who already have a lexicon in place (however, see Swingley & Aslin, 2007, for an eye-tracking study with newly learned words), which makes them unsuitable for early segmentation research. Second, the Conditioned Headturn (CHT) Procedure, in which infants are trained to turn to a puppet box for reinforcement each time they hear a target word, can also be used to test infants' extraction of words from fluent speech. In CHT studies on phoneme discrimination, target words or syllables were embedded in a list of other words, all spoken in isolation (Werker, Polka, & Pegg, 1997), but more recently, infants have been trained to respond to target words embedded in utterances (Dietrich, 2006; Gout, Christophe, & Morgan, 2004), and Gout et al. have claimed that CHT provides a more sensitive measure of word segmentation capabilities than HPP.

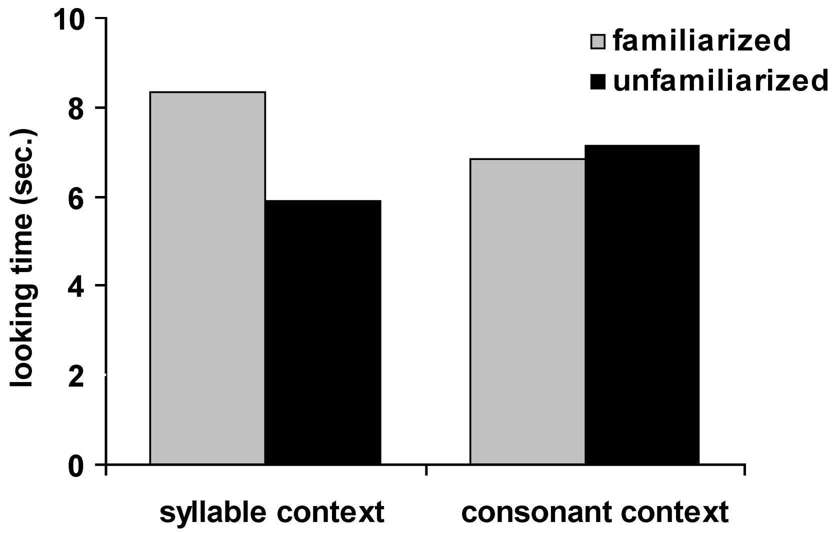

Figure 3: A summary of the results of Johnson, Jusczyk, Cutler & Norris, 2003. *The bar graph shows the looking times to familiar vs. unfamiliar words embedded in longer words, such as win embedded in window (syllable context) or wind (consonant context).*

Although the dependent measure in CHT is usually not the speed of initiating a headturn but the probability of making one, this method almost approaches an online measure, and it clearly has the potential to provide a useful convergent measure of early word segmentation. But CHT has a notoriously high dropout rate, and it typically requires two highly experienced experimenters to run the procedure. Given the skills needed to run CHT, procedural differences between laboratories could affect the reproducability of the results. Moreover, while HPP's familiarization phase is arguably a laboratory instantiation of natural parental repetitions, CHT's phase of training infants to attend to a specific word could be seen as less ecologically valid.

Online reflection of infant speech perception is, however, available from non-behavioral methods; in particular, electrophysiological methods have been used to study infant speech processing for over 30 years (Molfese, Freeman & Palermo, 1975). EEG and ERP are online measures with a high temporal resolution, which may be highly suitable for the study of word segmentation. In the next sections, these and other neuroimaging methods will be discussed.

EEG AND ERP

Electroencephalography

Electroencephalography (EEG) measures the electrical signals generated by the cortical, and to a lesser degree subcortical, areas of the brain. Cortical pyramidal cells firing in synchrony are for the most part responsible for the small voltage fluctuations that can be picked up by EEG measurements. Adult EEG typically has relatively low amplitudes (up to 100 μV) and is dominated by alpha (8 to 12 hz) and beta (12 to 30 hz) frequencies. Infant EEG contains higher amplitudes (up to 200 μV). In addition, it contains frequencies around 4-5 Hz, less common in adult EEG (see Figure 4), whereas alpha frequencies do not reach a mature level until after the first year of life.

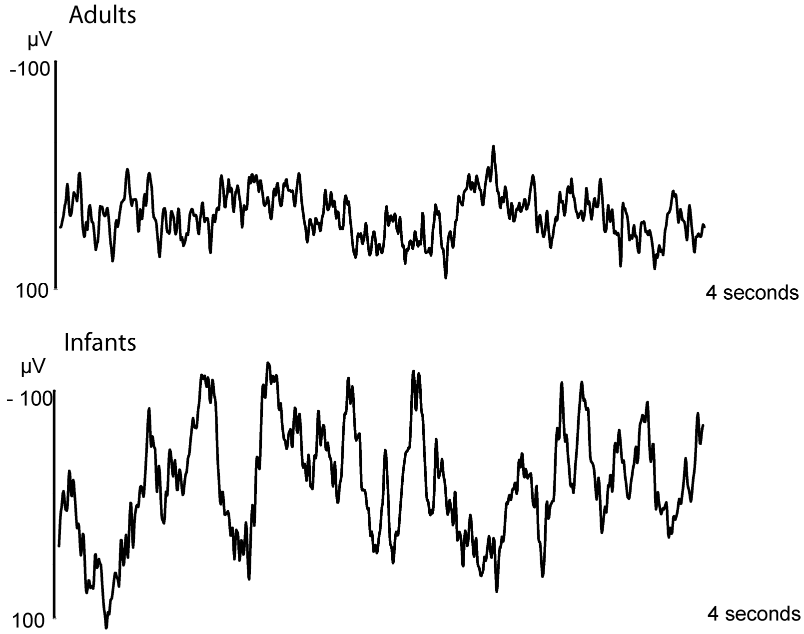

***Figure 4: Infant and adult EEG**. The infant's electroencephalogram (EEG; bottom) in general contains more slow frequencies and higher amplitudes than the adult EEG (top).*

Both physical and neural changes are responsible for these differences in background EEG. Signal conduction is affected by skull thickness (Grieve, Emerson, Isler, & Stark, 2004) and closing of the fontanels (Flemming, et al., 2005). Neural changes continuing well after birth such as synaptic and dendritic growth, myelination, and cortical folding (Uylings, 2006) most likely also affect the EEG signal. Therefore, direct comparisons between infant and adult EEG should be carefully interpreted, and cognitive, neural and physical differences have to be taken into account. Nevertheless, there are major advantages to using EEG as a tool to study cognitive development. The most important advantage is the high temporal resolution which gives insight into the speed and order of cognitive processes at a millisecond level. This makes it a very useful tool to study language development, and in particular sentence processing, since the temporal nature of speech can be taken into account. In addition, EEG can be used relatively easy with difficult subject groups such as young children, because it is an easy noninvasive procedure and does not require the subject to perform an overt task. The use of so-called EEG caps, i.e. caps containing a number of electrodes in fixed positions, has further increased the utility of EEG with infants.

Event Related Brain Potentials

Event Related Brain Potentials (ERPs) can be extracted from the EEG signal and give insight into the cognitive processes addressed in the experimental design. Figure 5 illustrates an EEG measurement and the extraction of ERPs from the EEG signal. In a conventional ERP session, a cap containing a number of electrodes is placed on the subject's head to measure the EEG signal, as well as eye electrodes to measure eye movements. Eye electrodes are placed at supra- and suborbital positions to measure vertical eye movements, and at right and left canthal positions to measure horizontal eye movements. A reference electrode is usually placed on a relatively neutral position, such as the nose or mastoid bone. The skin under the electrodes is cleaned with alcohol and abrasive paste to reduce skin impedance, after which the electrodes are filled with an electrolyte paste. This paste conducts the signal from the skull to the electrode. The electrodes transport the signal to an amplifier, which in turn transports the signal to a computer. In a typical cognitive ERP experiment, stimuli are presented to the participant during continuous EEG recording. A marker, usually time-locked to the onset of stimulus presentation (but sometimes also to the offset of the stimulus, or to the participant's response) is linked to the EEG signal. Offline, the EEG signals to different stimulus types (i.e., conditions) are extracted and visually inspected for artifact. Possible sources of artifact are eye movements, blinks, muscle activity in the face or neck of the participant or excessive motor activity. Trials with artifact are usually removed or corrected with automatic correction procedures. After artifact correction, the trials are averaged for each condition and for each subject, thus calculating subject averages. Unlike the EEG signal, ERPs only have very small amplitudes (in general less than 1 to 10 μV). Averaging over a number of occurrences of the same stimulus type reduces random brain activity and reveals the ERP components related to a particular cognitive event. Due to high individual variation, the grand average waveforms are usually reported and not the subject averages. The grand averages are calculated by averaging the subject averages per condition.

The ERP consists of a series of positive and negative peaks, or *components*. These ERP components are usually described by their peak latency and polarity. For example, the N400 is the name given to a component with a negative polarity (N) and a peak latency at about 400 ms after stimulus onset. The P300 is a component with a positive polarity (P) and a peak latency of 300 ms. The components can also be described in an ordinal manner. For example, the N1 refers to the first negative peak and P2 refers to the second positive peak after stimulus onset. In general, the early components of the ERP are referred to as exogenous components. These occur in the first 100 ms after stimulus onset and are mainly evoked by physical characteristics of the stimulus in the primary sensory pathways (0-10 ms) and thalamic areas (10-100 ms). Endogenous components, which occur after 100 ms, are for the most part responsive to cognitive processes and have their origin in the cortical areas of the brain. In language research, these endogenous components are of interest as they can reflect cognitive processing as a response to linguistic stimuli.

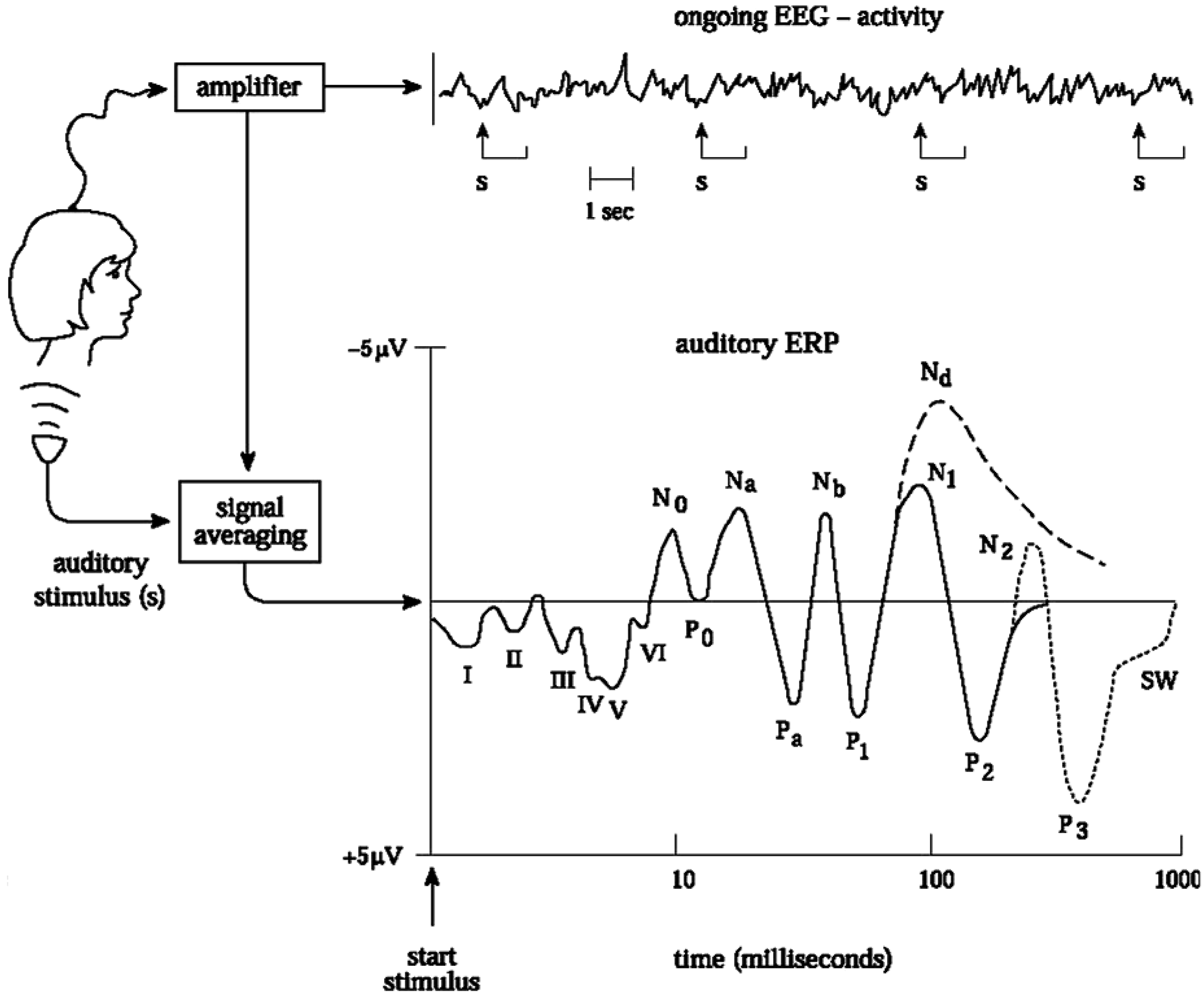

Figure 5: ERP measurement. *Graphic representation of an auditory ERP experiment. The ERP (bottom) is in general too small to be detected in ongoing EEG (top), and it requires averaging over a large number of stimulus presentations to achieve an adequate stimulus-to-noise ratio. It is assumed that by averaging the EEG, all randomly distributed activity is removed and only activity related to the previous cognitive event is left over in the ERP. The auditory ERP shown in this figure has a logarithmic time scale (bottom). This allows us to see the exogenous (I-VI, N0-Nb, and P1, N1, and P2) as well as the endogenous (Nd, N2, P300 and slow wave) ERP components. (After Hillyard & Kutas, 1983.)*

ERP components to different types of stimuli can be compared in terms of amplitude, voltage, distribution over the head, and onset latency, and give insight in differences in processing. Differences in amplitude are referred to as *ERP effects,* and reflect the relative amount of processing needed for certain stimuli types. Changes in the distribution of voltages over the head indicate that (partly) different underlying processes (and generators) are involved. The onset latency of an ERP component provides a measure of the speed with which the different stimulus types are processed at a millisecond level. This high temporal resolution is considered one of the most important strengths of ERP research, in addition to the non-invasiveness of the procedure and the possibility to omit an overt task.

A major weakness of ERP is the low spatial resolution. As explained above, differences in voltage distribution over the head suggest that (partly) different generators are involved in different conditions. Thus, it is possible to conclude *that* different generators are involved but it is extremely difficult to establish *which* generators. It requires a high number of electrodes (as many as 128 to 256) placed with an even distribution over the head, and quantitative techniques such as dipole modeling to make even a rough estimate of the underlying generators (Grieve, et al., 2004). Knowing the distribution of voltages over the head, as we do with EEG, does not provide enough information onwhich to base a precise estimate of the sources involved. Any number of dipoles with any combination of orientations can cause the voltage distribution observed. This is called the *inverse problem* of EEG (for a detailed description of EEG and the inverse problem, see Luck, 2005). Other neuroimaging tools, such as Magnetic Resonance Imaging (MRI), provide a much better spatial resolution but a very low temporal resolution.

ERP and language studies

Many laboratories use ERPs to investigate language processing, and quite a few have now turned to the use of ERPs to study language development. In adults, ERPs have been used for a considerable number of years as a measure of language processing. In this time, several ERP components have been well described. For example, the N400 has been shown to be related to semantic information processing (Federmeier & Kutas, 1999; Holcomb & Neville, 1991; Kutas & Federmeier, 2000). Grammatical information processing has been shown to be reflected by the (Early) Left Anterior Negativity (Friederici, Hahne, & Mecklinger, 1996) and the SPS/P600 (Hagoort, Brown, & Osterhout, 1999). However, so far only a few ERP studies have been done on adult word segmentation. Sanders and Neville (2003a; 2003b) studied N1 modulation as a measure of word segmentation in both native and nonnative listeners. They found a larger N1 component to word initial syllables as compared to word medial syllables in the native English speakers; the nonnative listeners did not show N1 modulation. These results suggest altered word segmentation skills in nonnative listeners with knowledge of English. Nazzi, et al. (in press) found similar N1 modulation as Sanders and Neville (2003a) in French native listeners. In chapter 5 of this thesis an ERP study with Dutch and English listeners to Dutch is described.

Although we as yet know relatively little about ERP components in infants, this field of research is developing rapidly (for recent reviews, see Friederici, 2005; Kuhl, 2004). Overall, it appears that ERP components common in adults are already present to some extent at a young age. However, these components do not seem to reach a mature level until the second decade of life. The development of the N1/P2 complex as a response to tones shows considerable changes in amplitude and does not reach its mature level until about 14-16 years of age (Pasman, Rotteveel, Maassen, & Visco, 1999). The Mismatch Negativity (MMN) response is a measure of perceptual change detection. It is a useful tool to study phoneme perception and discrimination (e.g., Dehaene-Lambertz & Pena, 2001; Pang, et al., 1998), and word discrimination (Weber, Hahne, Friedrich, & Friederici, 2004), and can be detected from a very early age (e.g., Cheour et al., 1998). Developmental changes do occur however (Cheour, Leppänen, & Kraus, 2000). Onset and peak latency reduces with age during infancy and childhood. In addition, peak amplitude of the MMN increases in the first year of life. Scalp distribution of the MMN seems to be broader and more central in infants than in adults. Also, there is large individual variation in infants and differences in polarity have been reported (Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). For example, in a phonetic discrimination study, Rivera-Gaxiola et al. (2005) showed a MMN in seven-month-olds, with an expected negative polarity in some of the infants but with a positive polarity in others. A follow-up study showed this polarity difference to be predictive of non-native discrimination skills at eleven months of age. However, in a MMN study by Weber et al. (2004), individual variation and a low signal-to-noise ratio were due to excess slow wave activity common in young infants. Offline high-pass filtering at 1 Hz revealed the MMN responses. Thus, individual variation arises not only from differences in cognitive development but also from physical characteristics of the EEG signal and low signal-to-noise ratios.

Further ERP methods have been developed to study other aspects of language development. Mills et al. studied word recognition in 14- and 20-month-olds using a word list paradigm (Mills, et al., 2004). They found a negative response in the 200-400 ms time window to known versus unknown words. This response had a broad distribution in 14-month-olds but a left temporal and parietal distribution in 20-month-olds. The same paradigm was used to study phonetic representations in the early lexicon. In contrast to the findings of Swingley and Aslin (2000), but in line with those of Werker et al. (2002), a differential ERP response was found to known words and highly similar phonetic foils in 20-month-olds, but not in 14-month-olds. Mills et al. argued that these results show that infants indeed do not show detailed phonetic representations of their first words. However, differences between these three studies (Mills et al., 2004; Swingley & Aslin, 2000; Werker et al., 2002) may be responsible for the differences in results. For example, Mills et al. used word lists, whereas Swingley and Aslin used a word-picture matching paradigm with highly familiar words, and Werker et al. used a similar paradigm with novel words and objects. Such differences tap different levels of processing and require different cognitive skills.

Friedrich and Friederici (2004; 2005) used an ERP version of the word-picture matching task to study the N400 component as a representation of word meaning. They observed a N400-like semantic incongruity effect in 14- and 19-month-olds to known words incongruent with a picture of a familiar object. Holcomb, Coffey and Neville (1992) performed a study on the N400 in an auditory and visual sentence processing task in the age range of 5 to 26 years. They observed contextual priming effects (including the N400) in all age groups, but also considerable differences in distribution of these effects, and a reduction in amplitude and latency of different ERP components. The differences in distribution, in general until about age 13 to 16, may point to the involvement of different neural systems at different ages, but may also be due to brain maturation. The changes in amplitude and latency, probably due to changes in brain maturation, had a linear character and occurred from five to about 16 years of age. Thus, even though a N400-like effect can be observed as young as 14 months of age, considerable changes do occur throughout childhood.

The only ERP studies so far on the development of word segmentation from continuous speech in infants are by the author of this thesis. In the experimental chapters 2, 3, and 4 of this thesis the results of these ERP studies will be discussed.

Other neuroimaging techniques

Although not very common yet, several other neuroimaging techniques, including different types of EEG analyses, are now also used to study infant cognitive development. Quantitative EEG analysis (i.e. the study of frequencies present in the EEG signal) has been used to study visual attention in 8- to 11-month-olds. A sharp increase in frontal theta (4-8 Hz) activity can be seen during internally controlled attention (Orekhova, Stroganova, & Posikera, 1999); alpha (8-12 Hz) synchronization over the posterior cortex was proposed to be involved in maintaining attention (Orekhova, Stroganova, & Posikera, 2001). Source localization in infants looking at novel visual stimuli was studied with high impedance amplifiers and a 124 channel EEG system (Reynolds & Richards, 2005). A large central negativity (Nc) was found and localized in the prefrontal cortex and the anterior cingulate cortex. However, as stated before, other neuroimaging techniques are much better suited to the study of source localization. In addition, Grieve et al. (2004) showed a higher error in spatial distribution estimates in infant EEG, mostly because infant brain areas are closer together, making source localization even more difficult.

Neuroimaging measures other than EEG have also been used with infants. Magnetoencephalography (MEG) measures the magnetic fields produced by the electrical activity in the brain and is especially useful for source localization. This relatively new neuroimaging tool has been used to study auditory discrimination in 6- and 12-month-old infants using an adult-size MEG system (Cheour, et al., 2004). A large variability in the MEG of the infants was found and additional research with infant size MEG systems is needed to establish the value of MEG in developmental studies. Testing infants with adult MEG systems introduces a large amount of noise to the data due to the distance between the sensors and the head. The use of infant MEG systems, with head coils proportionate to the size of the infant's head may solve this problem. In addition, movement of the head also causes a considerable amount of artifact. Smaller systems with infant size seats or beds may reduce this form of artifact and make MEG more usable with young children.

Optical Topography (OT) is a new technique that uses near-infrared light to measure changes in hemoglobin levels and blood volume in the brain. It was used to study lateralization of language processing in infants (Pena, et al., 2003). A left hemisphere dominance was found for speech stimuli (as compared to reversed speech) at two to five days after birth, using a 24-channel topography device. OT can be a valuable tool for localization studies, especially since functional Magnetic Resonance Imaging (fMRI) studies are usually not possible or ethically approved of in healthy infants. Another big advantage over fMRI is that OT operates silently. A limitation, however, is that the device cannot measure sources that are located at greater depth than three centimeters. This limits the usability of the technique to the surface of the brain, since deeper lying areas in the brain can not be represented. Very young infants, however, have thinner skulls and smaller heads than adults, so that a relatively large part of the brain can be studied.

With fMRI, blood oxygenation levels of the brain can be measured during task performance with a high spatial resolution. However, this method is difficult to use with young infants since it requires participants to be very still. In addition, the scanner produces a strong magnetic field, makes a lot of noise and is quite intimidating to participants. Therefore, so far only a few fMRI studies have been done with young infants (for a recent overview of fMRI studies on speech processing in infants, see Dehaene-Lambertz, et al., 2006). For example, a study with awake three-month-olds showed left hemisphere dominance for auditory perception (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002). In particular, the left angular gyrus showed a stronger activation to forward speech than to backward speech. In a recent study, brain activation to sentences showed a network of perisylvian areas in three-month-old infants (Dehaene-Lambertz, et al., 2006). This pattern of activation is comparable to that of adults listening to speech. In addition, repetition of a sentence resulted in a stronger response in Broca's area. In adults, this area is related to speech production, but also to comprehension and memory. The response to repetition in three-month-olds in this area may represent linguistic memory in infants. These exciting results show an important role for fMRI in language development in the future, despite the fact that it is a difficult technique to use with young infants.

Thus, neuroimaging techniques other than ERP are becoming increasingly popular as tools to study early cognitive development, but still suffer from some limitation. MEG is an expensive technique and infant-size MEG systems are not widely used yet. OT and MRI have high spatial resolution but low temporal resolution. In addition, it is extremely difficult to get ethical approval for MRI studies with children in the Netherlands. ERP, however, is a technique that can be used relatively easily with infants. It is a non-invasive technique that does not require an overt task. Moreover, ERP provides the high temporal resolution that can not be achieved with behavioral methods, but is inevitable for studies of online sentence processing. Thus, the development of word segmentation seems best tackled using ERP. In this thesis, the first ERP studies on infant word segmentation from continuous speech are presented, providing new insight into this important step in language development.

ISSUES AND OUTLINE OF THIS THESIS

In the remainder of this thesis, four experimental chapters are presented. In chapter 2, the first ERP study on word segmentation from continuous speech in infants is presented. In the study described in chapter 3, the role of stressed syllables in segmentation of words with an initial weak syllable was studied in ten-month-olds. Chapter 4 deals with word segmentation in seven-month-old infants. An ERP and a behavioral study were performed to gain deeper insight into early word segmentation in Dutch infants, and the way it is reflected by each type of task. In chapter 5, word segmentation in both native and foreign adult listeners to Dutch is compared. Finally, in chapter 6 the results and conclusions of the experimental chapters are summarized.

REFERENCES

Aslin, R.N., Jusczyk, P.W., & Pisoni, D.B. (1998). Speech and auditory processing during infancy: Constraints on and precursors to language. In D. Kuhn and R. Siegler (Eds.), *Handbook of Child Psychology*, (5th edition, vol. 5, pp. 147-198). New York: Wiley.

Bates, E., Thal, D., Finlay, B., & Clancy, B. (2002). Early language development and its neural correlates. In J. Segalowitz, and I. Rapin (Eds.), *Handbook of Neuropsychology, (*2nd edition, vol. 8, pp.S 109-176). Amsterdam: Elsevier.

Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

Bloomfield, L. (1933). *Language*. New York: Holt.

Bortfeld, H., Morgan, J.L, Golinkoff, R.M., & Rathbun, K. (2005). *Mommy* and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science 16*, 298-304.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in four-month-old infants from monolingual and bilingual environments. *Cognition 65*, 33-69.

Brent, M.R. (1999). Speech segmentation and word discovery: A computational perspective. *Trends in Cognitive Science 3*, 294-301.

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition 61*, 93-125.

Cairns, P., Shillcock, R., Chater, N., & Levy, J. (1997). Bootstrapping word boundaries: a bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology 33*, 111-153.

Chaney, C.F. (1989). I pledge a legiance tothe flag: Three studies in word segmentation. *Applied Psycholinguistics 10*, 261-281.

Chaney, C.F., & Estin, E.T. (1987). Metalinguistic awareness. Miniseminar presented at the American Speech-Language-Hearing Association, New Orleans.

Cheour, M., Alho, K., Ceponiené, R., Reinikainen, K., Sainio, K., Pohjavouri, M., Aaltonen, O., & Näätänen, R. (1998). Maturation of mismatch negativity in infants. *International Journal of Psychophysiology 29*, 217-226.

Cheour, M., Imada, T., Taulu, S., Ahonen, A., Salonen, J., & Kuhl, P. (2004). Magnetoencephalography is feasible for infant assessment of auditory discrimination. *Experimental Neurology 190*, S44-S51.

Cheour, M., Leppänen, P.H.T., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology, 111*, 4-16.

Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: a connectionist model. *Language and Cognitive Processes 13*, 221-268.

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science 1*, 215-219.

Chugani, H.T., Phelps, M.E., & Mazziotta, J.C. (1987). Positron emission tomography study of human brain functional development. *Annals of Neurology 22*, 487-497.

Clark, E.V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences 8*, 472-478.

Cole, R.A., & Jakimik, J. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.). *Strategies of Information Processing* (pp. 67-116). London, Academic Press.

Cutler, A. (2001). Listening to a second language through the ears of a first. *Interpreting 5*, 1-23.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language 31*, 218-236.

Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language 2*, 133-142.

Cutler, A., Demuth, K., & McQueen, J.M. (2002). Universality versus language-specificity in listening to running speech. *Psychological Science 13*, 258-262.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of experimental psychology. Human perception and performance 14*, 113-121.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science 298*, 2013-2015.

Dehaene-Lambertz, G., Hertz-Pannier, L., & Dubois, J. (2006). Nature and nurture in language acquisition: anatomical and functional brain imaging studies in infants. *Trends in Neurosciences 29*, 367-373.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., & Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences 103*, 14240-245.

Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport 12*, 3155-3158.

DeCasper, A.J., & Fifer, W.P. (1980). Of human bonding: Newborns prefer their mothers' voice. *Science 208*, 1174–1176.

DeCasper, A.J., & Spence, M.J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development 9*, 133–150.

Dietrich, C. (2006). The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. PhD Dissertation, Radboud University Nijmegen. *MPI Series in Psycholinguistics 40*.

Eimas, P.D., Siqueland, E.R., Jusczyk, P.W., & Vigorito, J. (1971). Speech perception in infants. *Science 171*, 303–306.

Elman, J.L. (1990). Finding structure in time. *Cognitive Science 14*, 179–211.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language 41*, 469-495.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development 8*, 181-195.

Fernald, A., Pinto, J.P., Swingley, D., Weinberg, A., & McRoberts, G.W. (1998). Rapid gains in speed of verbal processing by infants in the second year. *Psychological Science 9*, 228-231. (Reprinted in Tomasello and E. Bates (eds.)., *Language Development: The Essential Readings,* M. Blackwell, 2001.)

Flemming, L., Wang, Y.Z., Caprihan, A., Eiselt, M., Haueisen, J., & Okada, Y. (2005). Evaluation of the distortion of EEG signals caused by a hole in the skull mimicking the fontanel in the skull of human neonates. *Clinical Neurophysiology 116*, 1141-1152.

Friederici, A.D. (2005). Neurophysiological markers of early language acquisition: from syllables to sentences.*Trends in Cognitive Sciences 9*, 481-488.

Friederici, A.D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition 22,* 1219–1248.

Friedrich, M., & Friederici, A.D. (2005). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport 16*, 653-656.

Friedrich, M., & Friederici, A.D. (2004). N400-like semantic incongruity effect in 19- month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience 16*, 1465-1477.

Ganger, J., & Brent, M.R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology 40*, 621-632.

Gaskell, M.G., & Marslen-Wilson, W.D. (1997). Integrating form and meaning: a distributed model of speech perception, *Language and Cognitive Processes 12*, 613-656.

Gield, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., & Rapoport, J.L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience 2*, 861-863.

Gout, A., Christophe, A., & Morgan, J.L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language 51*, 548-567.

Grieve, P.G., Emerson, R.G., Isler, J.R., & Stark, R.I. (2004). Quantitative analysis of spatial sampling error in the infant and adult electroencephalogram. *Neuroimage 21*, 1260-1274.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C.M. Brown and P. Hagoort (Eds.), *The neurocognition of language*, (pp. 273-316). New York: Oxford University Press.

Hillyard, S.A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34*, 33-61

Holcomb, P.J., Coffey, S.A., & Neville, H.J. (1992). Visual and auditory sentence processing: a developmental analysis using event-related brain potentials. *Developmental Neuropsychology, 8*, 203-241.

Holcomb, P.J., & Neville, H.J. (1991). Natural speech processing: an analysis using event-related brain potentials. *Psychobiology, 19*, 286-300.

Holden, M.H., & MacGinitie, W.H. (1972). Children's conception of word boundaries in speech and print. *Journal of Educational Psychology 63*, 551-557.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R.M. (2000). Breaking the language barrier:An emergentist coalition model for the origins of word learning.

Monographs of the Society for Research in Child Development 65 (3, Serial No. 262).

Houston, D.M., & Horn, D.L. (submitted). Assessing speech discrimination in individual infants.

Huttenlocher, J. (1964). Children's language: word-phrase relationship. *Science 143*, 264-265.

Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language 44*, 548-567.

Johnson, E.K., Jusczyk, P.W., Cutler, A., & Norris, D. (2003). Lexical viability constraints on speech segmentation by infants. *Cognitive Psychology 46*, 65-97.

Jusczyk, P.W. (1999). How infants begin to extract words from fluent speech. *Trends in Cognitive Science 3*, 323-328.

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P.W. (1985). The high-amplitude sucking technique as a methodological tool in speech perception research. In G. Gottlieb and N.A. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life: A methodological overview* (pp. 195–222). Norwood, NJ: Ablex.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology 29*, 1–23.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology 39*, 159–207.

Kemler Nelson, D.G., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A., & Gerken, L.A. (1995). The Headturn Preference Procedure for testing auditory perception. *Infant Behavior and Development 18*, 111-116.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience 5*, 831-843.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom B. (1992). Linguistic experience alters phonetic perception in infants by six months of age. *Science 255*, 606-608.

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences 4*, 463-470.

Luck, S. (2005). *An introduction to the Event-Related Potential technique*. Cambridge, MA: MIT Press.

Mandel, D.R., Jusczyk, P.W., & Pisoni, D.B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science 6*, 314–317.

Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology 10*, 29-63.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology 38*, 465-494.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology-General 134*, 477-500.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology 18*, 1-86.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language 39*, 21-46.

Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience 16*, 1452-1464.

Molfese, D.L., Freeman, R.B. Jr., & Palermo. D.S. (1975). The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain and Language 2*, 356-368.

Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child development 66*, 911-936

Moon, C., Cooper, R.P., & Fifer, W.P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development 16*, 495–500.

Morgan, J.L. (1996). Prosody and the roots of parsing. *Language and Cognitive Processes 11*, 69-106.

Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development 66*, 911-936.

Mrzljak, L., Uylings, H.B.M., Van Eden, C.G., & Judas, M. (1990). Neuronal development in human prefrontal cortex in prenatal and postnatal stages. *Progress in Brain Research 85*, 185-222.

Myers, J., Jusczyk, P.W., Kemler Nelson, J., Charles-Luce, J., Woodward, A.L., & Hirsch-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language 23*, 1-30.

Nazzi, T., Dilley, L.C., Jusczyk, A.M., Shattuck-Hufnagel, S., & Jusczyk, P.W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech 48, 279-298*.

Nazzi, T., Iakimova, G., Bertoncini, J., & de Schonen, S. (in press). *Behavioral and electrophysiological exploration of early word segmentation in French: Distinguishing the syllabic and lexical levels*. In A. Friederici and G. Thierry (Eds.), *Trends in Language Acquisition Research, Amsterdam: John Benjamins Publishing.*

Nazzi, T., Jusczyk, P.W., & Johnson, E.K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language 43*, 1-19.

Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology 41*, 352-362.

Newman, R., Bernstein Ratner, N., Jusczyk, A.M., Jusczyk, P.W., & Dow, K.A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology 42*, 643-655.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition 52*, 189-234.

Norris, D., McQueen, J.M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition 21*, 1209-1228.

Norris, D., McQueen, J.M., Cutler, A., & Butterfield, S. (1997). The Possible-Word Constraint in the segmentation of continuous speech. *Cognitive Psychology 34*, 191-243.

Orekhova, E. V., Stroganova, T. A., & Posikera, I. N. (2001). Alpha activity as an index of cortical inhibition during sustained internally controlled attention in infants. *Clinical Neurophysiology 112*, 740-749.

Orekhova, E. V., Stroganova, T. A., & Posikera, I. N. (1999). Theta synchronization during sustained anticipatory attention in infants over the second half of the first year of life. *International Journal of Psychophysiology 32*, 151-172.

Pang, E.W., Edmonds, G.E., Desjardin, R., Khan, S.C., Trainor, L.J., & Taylor, M.J. (1998). Mismatch negativity to speech stimuli in 8-month-old infants and adults. *International Journal of Psychophysiology*, *29***,** 227–236.

Pasman, J.W., Rotteveel, J.J., Maassen, B., & Visco, Y.M. (1999). The maturation of auditory cortical evoked responses between (preterm) birth and 14 years of age. *European Journal of Paediatric Neurology 3*, 79-82.

Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 11702-11705.

Peperkamp, S. (2003). Phonological acquisition: Recent attainments and new challenges. *Language and Speech 46*, 87-113.

Pujol, J., Soriano-Mas, C., Ortiz, H., Sebastián-Gallés, N., Losilla, J.M., & Deus, J. (2006). Myelination of language-related areas in the developing brain. *Neurology 66*, 339-343.

Reynolds, G.D., & Richards. J.E. (2005). Familiarization, attention, and recognition memory in infancy: An Event-Related Potential and cortical source localization study. *Developmental Psychology 41*, 598-615.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P.K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science 8*, 162-172.

Saffran, J.R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition 81*, 149-169.

Sameroff, A.J. (1967). Nonnutritive sucking in newborns under visual and auditory stimulation. *Child Development 38*, 443–452.

Sanders, L.D., & Neville, H.J. (2003a). An ERP study of continuous speech processing I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research 15*, 228-240.

Sanders, L.D., & Neville, H.J. (2003b). An ERP study of continuous speech processing II. Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research 15*, 214-227.

Seidl, A., & Johnson, E.K. (forthcoming). Utterance boundaries and infants' extraction of words from fluent speech.

Soderstrom, M., Kemler Nelson, D.G., & Jusczyk, P.W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior and Development 28*, 87-94.

Swingley, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition 76*, 147-166.

Swingley, D., & Aslin, R.N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology 54*, 99-132.

Swingley, D., Pinto, J.P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition 71*, 73-108.

Toro, R., & Burnod, Y. (2005). A morphogenetic model for the development of cortical convolutions. *Cerebral Cortex 15*, 1900-1913.

Tincoff, R., & Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science 10*, 172–175.

Tunmer, W.E., Bowey, J.A., & Grieve, R. (1983). The development of young children’s awareness of words as a unit of spoken language. *Journal of Psycholinguistic Research 12*, 567-594.

Uylings, H.B.M. (2006). Development of the human cortex and the concept of 'critical' or 'sensitive' periods. *Language Learning 56* (s1), 59-60.

Van de Weijer, J. (1998). Language input for word discovery. PhD Dissertation, Radboud University Nijmegen. *MPI Series in Psycholinguistics 9*.

Weber, C., Hahne, A., & Friedrich, M., & Friederici, A.D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence.*Cognitive Brain Research, 18*, 149-161.

Werker, J.F. (2003). Baby steps to learning language. *Journal of Pediatrics* 143, S62-S69.

Werker, J.F., Fennell, C.T., Corcoran, K.M., & Stager, C.L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy 3*, 1-30

Werker, J.F., Polka, L., & Pegg, J.E. (1997). The Conditioned Head Turn Procedure as a method for testing infant speech perception. *Early Development and Parenting* 6, 171-178.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development 7*, 49–63.

Werker, J.F., & Tees, R.C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology 50*, 509–535.

Wolff, J.G. (1977). The discovery of segments in natural language. *British Journal of Psychology 68*, 97-106

Woodward, J.Z., & Aslin, R.N. (1990). Segmentation cues in maternal speech to infants. Paper presented at the meeting of the International Conference of Infant Studies, Montreal, Quebec, Canada.

Electrophysiological evidence of prelinguistic infants' word recognition in continuous speech

Chapter 2

This chapter is a slightly adjusted version of the paper Kooijman, V., Hagoort, P., & Cutler, A., 2005. Electrophysiological evidence of prelinguistic infants' word recognition in continuous speech. Cognitive Brain Research, 24: 109-116.

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, pre-linguistic 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.

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. Figure 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 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 (Van de Weijer, 1999), 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 Figure 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, i.e. 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 attempt 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 (Jusczyk, 1999). 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 (Houston, Jusczyk, Kuijpers, Coolen & Cutler, 2000; Jusczyk, 1999; Jusczyk & Aslin, 1995; Jusczyk, Houston & Newsome, 1999; Kuijpers, Coolen, Houston & Cutler, 1998). 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.

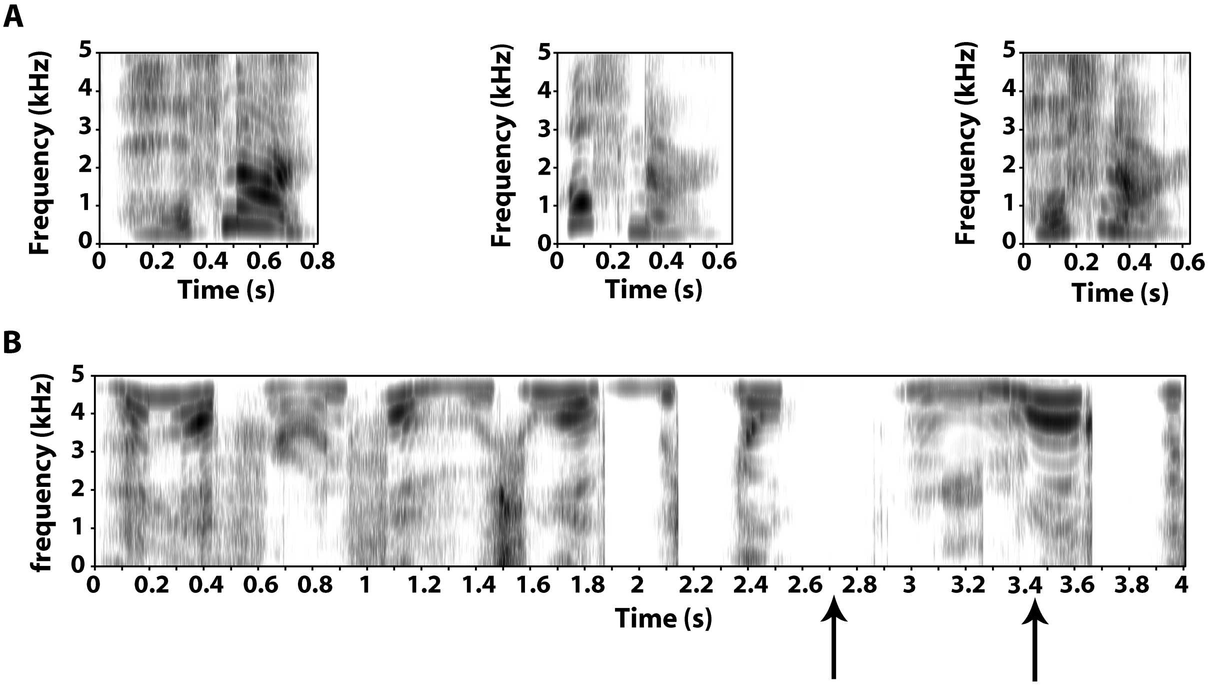

Figure 1: Spectrograms. *Figure 1A: Spectrograms of three utterances of the word hofnar ('court jester') spoken in isolation in an animated infant-directed manner; Figure 1B: Spectrogram of the sentence De koning hoort de boze hofnar vallen ('The king hears the nasty court jester falling'), spoken in 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 Figure A that these three utterances differ considerably, both in duration and in distribution of energy across the frequency spectrum. It can be seen from Figure 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.*

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 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 (Cheour, Leppänen & Kraus, 2000), and studies have also been conducted on discrimination of (isolated) pseudowords (e.g. in 4- and 5-month-old infants: see Weber, Hahne, Friedrich & Friederici, 2003) and (isolated) words (4-7 year-old children: see Korpilahti, Krause, Holopainen & Heikki Lang, 2001). However, this type of paradigm is less optimal for answering the current research question, for which more complex stimuli, e.g. 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 (Rugg & Doyle, 1994), 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 pre-linguistic 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 (Baayen, Piepenbrock & Van Rijn, 1993; Cutler & Carter, 1987), and the head turn preference response has been consistently observed for such words in both languages (Houston et al., 2000; Jusczyk et al., 1999; Kuijpers et al., 1998). The ten 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 Figure 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).

Table 1: An example of an experimental block (with literal English translation between brackets).

Familiarization phase: 10x *python*

Test phase:

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

MATERIALS AND METHODS

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 was 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.

Materials

Forty low frequency, two-syllable nouns (from here onwards: target words; see Table 2 for a list of all target words, and *Appendix 1A* for all materials) with a strong-weak stress pattern (that is, stress on the first syllable) were selected from the CELEX Dutch lexical database (Baayen et al., 1993). 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.

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

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 synchronised 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 minutes, 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.

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) (American Electroencephalographic Society, 1994). 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 band pass of 0.1-30 Hz and a sample rate of 200 Hz. Impedances were below 10 kΩ 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 Mills, Coffey-Corina & Neville, 1993). 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.

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 mean waveform (Figure 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. Figure 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=.004). This Familiarity effect differed across quadrants (significant Familiarity x Quadrant interaction: $F_{1,27}$=6.34, p=.002). In separate analyses by quadrant, the Familiarity effect was significant in the left and right frontal quadrants ($F_{1,27}$=19.45, p<.001; $F_{1,27}$=10.84, p=.003, respectively). Figure 2c illustrates the distribution of this effect in an isovoltage plot. (Also see *Supporting Table 3, Appendix 2A*). Next, we determined the exact onset of the Familiarity effect by examining the difference waveform (word position 9/10 – 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.; see also Van Berkum, Brown, & Hagoort, 1999). Significance (p<.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; $p<.05$)). 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. Figure 3a shows grand mean 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 over the right hemisphere (Figure 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, $p>.05$), but we did find a significant interaction for Familiarity by Hemisphere ($F_{1,27}=5.01$, $p=.034$). In separate analyses by hemisphere, we found a significant Familiarity effect over the left hemisphere ($F_{1,27}=4.84$, $p=.037$), but not over the right ($p>.05$). (Also see *Supporting Table 4, Appendix 2A.*) 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 $p<.05$ 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, $p<.05$). This response to the Test materials is, as in the Familiarization phase, positive-going, but begins later.

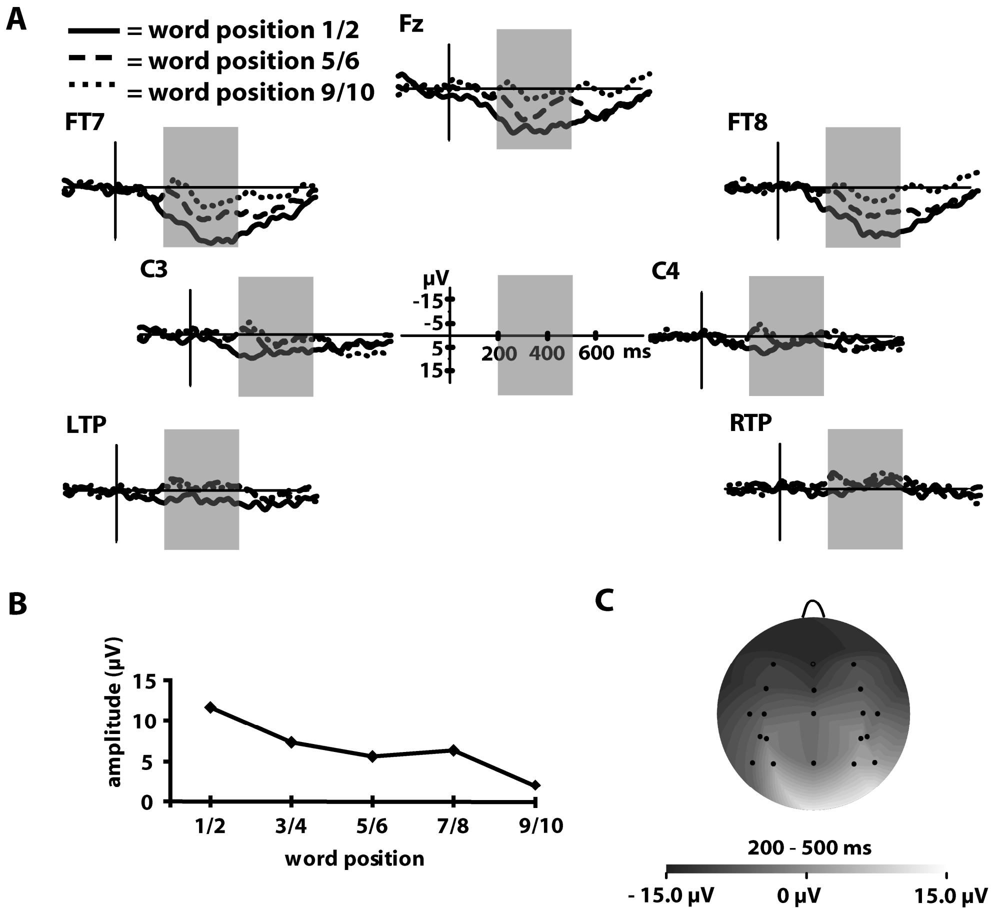

Figure 2: Results of the Familiarization phase. *Figure 2A: The grand mean waveforms to word position 1/2, 5/6 and 9/10 at seven representative electrode positions Fz, FT7, FT8, C3, C4, LTP and RTP; negativity is plotted upwards. The grey area indicates the time window from 200-500 ms from word onset.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

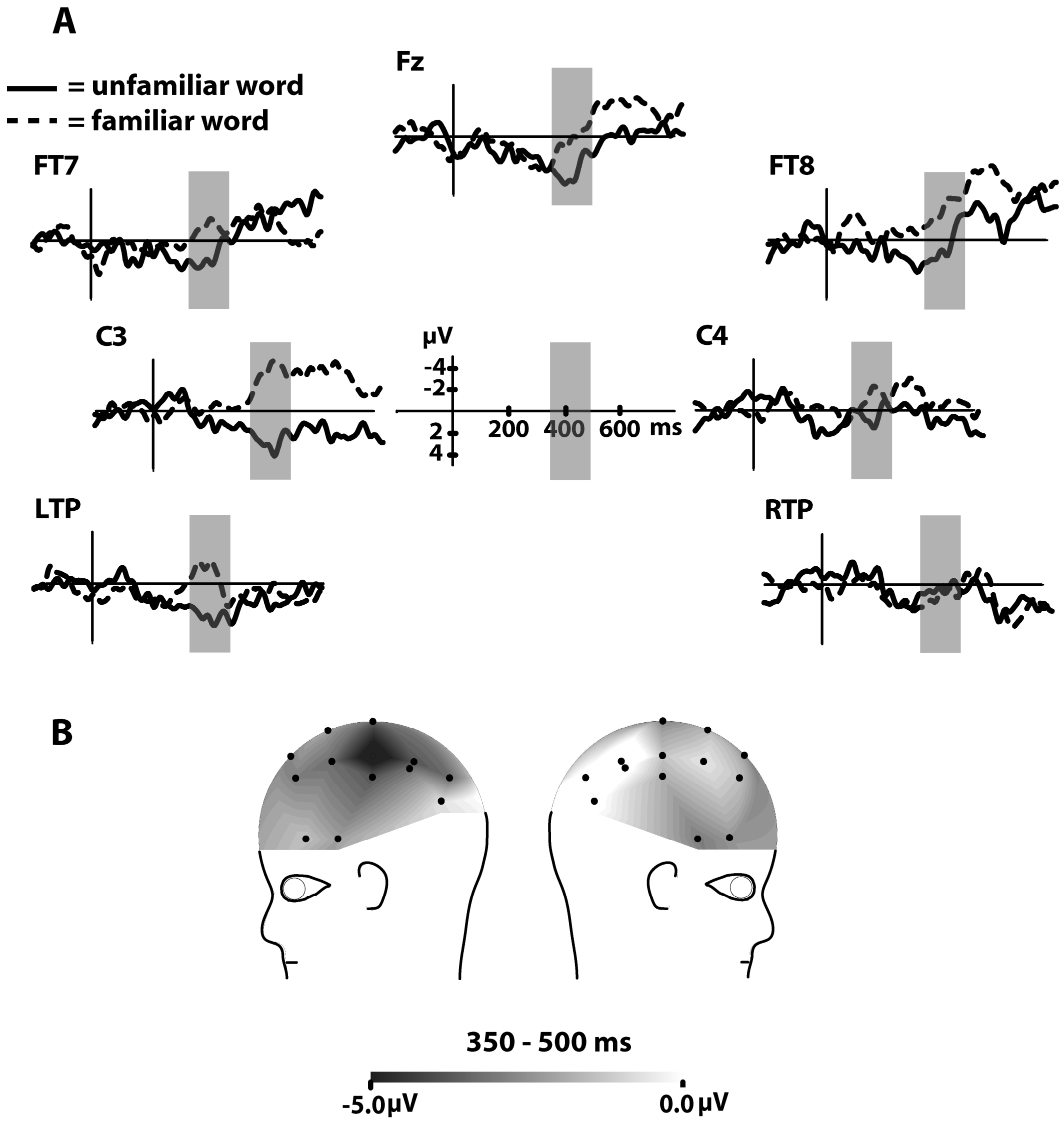

Figure 3: Results of the Test phase. *Figure 3A: The grand mean waveforms to the unfamiliar words and the familiar words in the sentences at seven representative electrode positions Fz, FT7, FT8, C3, C4, LTP and RTP; negativity is plotted upwards. The gray area indicates the time window from 350-500 ms from word onset. Figure 3B: 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 to the unfamiliar words from the ERP to the familiar words in the 350-500 ms latency range.*

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 Figure 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 to 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, i.e. 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 and collaborators (Houston et al., 2000; Jusczyk, 1999; Jusczyk et al., 1995; Jusczyk et al., 1999; Kuijpers et al., 1998) 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 (Heinze, Münte & Mangun, 1994; Mills et al., 1993; Rugg, Doyle, & Holdstock, 1994) 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 (Jusczyk et al., 1999), 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 (Van de Weijer, 1999), 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.

REFERENCES

American Electroencephalographic Society (1994). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111-113.

Baayen, R.H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

van Berkum, J.A., Brown, C.M., & Hagoort, P. (1999). Early referential context effects in sentence processing: evidence from event-related brain potentials. *Journal of Memory and Language, 41*, 147-182.

Cheour, M., Leppänen, P.H.T., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology, 111*, 4-16.

Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133-142.

Heinze, H.J., Münte, T.F., & Mangun, G.R. (Eds.). (1994). *Cognitive Electrophysiology*. Boston: Birkhäuser.

Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review, 3*, 504-509.

Jusczyk, P.W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*, 323-328.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1-23.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159-207.

Korpilahti, P., Krause, C.M., Holopainen, I. & Heikki Lang, A. (2001). Early and late Mismatch Negativity elicited by words and speech-like stimuli in children. *Brain and Language, 76,* 332-339.

Kuijpers, C., Coolen, R., Houston, D., & Cutler, A. (1998). Using the Head-Turning technique to explore cross-linguistic performance differences. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Advances in Infancy Research* (Vol. 12, pp. 205-220). London: Ablex.

Mills, D.L., Coffey-Corina, S.A., & Neville, H.J. (1993). Language acquisition and cerebral specialization in 20-month-old infants. *Journal of Cognitive Neuroscience, 5*, 317-334.

Rugg, M. D., & Doyle, M. C. (1994). 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*. Boston: Birkhäuser.

Rugg, M.D., Doyle, M.C., & Holdstock, J.S. (1994). Modulation of event-related brain potentials by word repetition: effects of local context. *Psychophysiology, 31*, 447-459.

Van de Weijer, J. (1999). Language input for word discovery [doctoral dissertation]. *MPI series in Psycholinguistics, 9*. Nijmegen: MPI.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A.D. (2003). Discrimination of word stress in early infant perception: electrophysiological evidence. *Cognitive Brain Research, 18,* 149-161.

Language-specific prosodic structure in early word-segmentation: ERP evidence from Dutch ten-month-olds

Chapter 3

This chapter is a slightly adjusted version of the paper Kooijman, V., Hagoort, P., & Cutler, A., under revision. Language-specific prosodic structure in early word- segmentation: ERP evidence from Dutch ten-month-olds. Infancy.

Recognizing word boundaries in continuous speech requires detailed knowledge of the native language. In the first year of life, infants acquire considerable word segmentation abilities. Infants at this early stage in word segmentation rely to a large extent on the metrical stress pattern of their native language, at least in stress-based languages. Segmentation of strong-weak words develops rapidly between seven and ten months of age. Nevertheless, trochaic languages contain not only strong-weak words but also words with a weak-strong stress pattern. In this paper, we present electrophysiological evidence of the beginnings of weak-strong word segmentation in Dutch ten-month-olds. At this age, the ability to combine different cues for efficient word segmentation does not yet seem to be developed completely. We provide evidence that Dutch infants still largely rely on strong syllables, even for the segmentation of weak-strong words.

INTRODUCTION

Before their first birthday, infants learn to extract possible word forms from continuous speech. In other words, they learn how to detect word boundaries in spoken language before they speak their first words. This ability, which is referred to as word segmentation, is very important for language development, and, as has been shown recently, is predictive of language skills at a later age, such as vocabulary size at age two (Newman, Bernstein Ratner, Jusczyk, Jusczyk, & Dow, 2006).

In adults, word segmentation is well established; adults can make use of many cues based on the probabilities of their native language (e.g., prosodic cues: Cutler & Butterfield, 1992; phonotactics: McQueen, 1998), and the effect of these cues is intensified in the larger adult vocabulary (Norris, McQueen, & Cutler, 1995). Infants beginning to learn their native language, however, do not have a vocabulary available to them yet, and cannot rely on lexical information. In addition, word boundaries do not correlate with silent breaks in spoken language. Instead, infants have to rely solely on cues in the sound structure of their native language, such as phonotactic (possible phoneme order) and phonetic (properties of speech sounds) regularities, and prosodic cues (e.g., the metrical stress pattern of a language). Unfortunately, these cues have a probabilistic rather than an all-or-none characteristic (Kuhl, 2004). Therefore, infants not only have to discover these separate cues, but they also have to learn how to combine them in order to detect word boundaries efficiently. Thus, learning to segment words from speech is not as easy a task as it may seem. Nevertheless, by about ten months of age infants are quite proficient at word segmentation.

One of the available cues in speech is the metrical rhythm of the native language, to which infants are sensitive from very early on. From birth onwards, infants show recognition of metric rhythm, both in language (Nazzi, Bertoncini, & Mehler, 1998; Nazzi & Ramus, 2003) and in music (Bergeson & Trehub, 2006). Newborns can discriminate languages from different rhythmic classes

(Nazzi et al., 1998). By four months, infants can discriminate between their native language and other rhythmically similar languages (Bosch & Sebastián-Gallés, 1997). At five months of age, infants show discrimination of strong-weak from weak-strong stress patterns presented in isolation (Weber, Hahne, Friedrich, & Friederici, 2004).

In the second half of the first year of life, English-learning infants prefer to listen to strong-weak words over weak-strong words (Jusczyk, Cutler & Redanz, 1993); in this, they show a preference for the words that are more typical in their language. Infants acquiring trochaic stress-based languages (i.e., languages with a predominantly strong-weak stress pattern) can recognize words which conform to this pattern when they occur in the context of continuous speech (English bisyllabic words: Jusczyk, 1999; Juszcyk, Houston & Newsome, 1999; Mattys & Jusczyk, 2001; English trisyllabic words: Houston, Santelmann, & Jusczyk,2004; Dutch bisyllabic words: Kuijpers, Coolen, Houston, & Cutler, 1998; Kooijman, Hagoort, & Cutler, 2005; submitted). At about nine months of age, infants show further knowledge of the typical patterns of their native language by demonstrating sensitivity to phonotactic properties (Jusczyk, Luce, & Charles-Luce, 1994; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Myers, Jusczyk, Kemler Nelson, Charles-Luce, Woodward, & Hirsch-Pasek, 1996; Saffran, Aslin, & Newport, 1996).

The predominant metrical stress pattern of the language seems to be a particularly salient cue for infants beginning to learn word segmentation. In a study using the Headturn Preference Procedure (HPP), English-learning infants listened to bisyllabic words with a phonotactic order between the two syllables that indicated either a within-word boundary or a between-word boundary (Mattys, Jusczyk, Luce, & Morgan, 1999). They showed a preference for strong-weak words with a within-word boundary, but for weak-strong words with a between-word boundary. Thus, nine-month-olds perceive the strong syllable as word onset and prefer to listen to the phonotactic order that corresponds to this perception. Next, Mattys et al. pitted prosodic cues against phonotactic cues by presenting the infants with strong-weak words with a phonotactic between-word boundary, and with weak-strong words with a phonotactic within-word boundary. Now the infants showed a preference for the strong-weak words in spite of the conflicting phonotactic word boundary information in these words. In an artificial language study, Johnson and Jusczyk (2001) also pitted metrical stress patterns against statistical distribution of speech sounds. After familiarization with a speech stream, infants were tested on isolated words and part-words with conflicting prosodic and phonotactic information. The results showed that, at eight months, English-learning infants weight prosodic cues more heavily than phonotactic cues. Thus, these studies show that infants prefer prosody over phonotactics in case of conflicting information. This preference, however, may lead to missegmentation of other word types from continuous speech, for example, words with a weak-strong stress pattern.

Although strong syllables are very salient in trochaic stress-based languages, metrical stress, like other possible segmentation cues, also has a distributional nature. The languages not only contain strong-weak words, but also a considerable number of words with a weak-strong stress pattern. Cutler and Carter (1987) showed that in a corpus of English a quarter consisted of words with an initial weak syllable. Taking frequency of occurrence into account, this percentage reduced to 17%. In a natural speech sample of English only 10% of the lexical words had a weak initial syllable (Cutler & Carter, 1987). In a study on a corpus of Dutch, it was shown that most of the speech infants hear comes from multi-word utterances and not words in isolation (Van de Weijer, 1998). In addition, 97.2% of the lexical words directed at an infant started with a strong syllable; speech directed at an older child and adults contained 96.4% and 88.3% lexical words with an initial strong syllable respectively.

Thus, the percentage of words with an initial weak syllable is relatively low in English and Dutch. However, it still accounts for a considerable number of words. At some point infants have to learn to deal with these words to efficiently segment all words from speech. A few studies have addressed this issue. A HPP study by Johnson (2005) showed that 10.5-month-old infants' representations of iambic (i.e., weak-strong) words are fairly detailed after familiarization. Thus, they do not seem to rely on just the strong syllable for words presented in isolation, but have a representation of the whole word, including the initial weak syllable. Jusczyk, Houston and Newsome (1999) ran an impressive series of HPP experiments to study weak-strong word segmentation in infants. They showed that English-learning 7.5-month-old infants are able to segment strong-weak words from speech, but not weak-strong words. However, 10.5-month-olds did show the ability to segment weak-strong words from speech. In addition, after familiarization with only the strong syllables of the weak-strong words, for example, *tar* from *guitar*, they did not then show a preference for passages containing the whole weak-strong words, as the 7.5-month-old infants had. These results suggest that while 7.5-month-olds may be just segmenting the strong syllables from speech, the 10.5-month-olds do more than that. The authors concluded that at this later age, infants no longer rely solely on the stress pattern of their native language for word segmentation, but are able to combine multiple sources of information about likely word boundaries in speech, such as metrical stress and phonotactics (also see Morgan & Saffran, 1995).

Considering the high similarity between English and Dutch (both are trochaic stress-based languages), a comparable rate of development of word segmentation might be expected. However, English-learning infants show a behavioral preference for familiar strong-weak words in sentences at 7.5 months of age (Jusczyk, 1999; Mattys et al., 2001), whereas Dutch infants do not show a behavioral preference until nine months of age (Kuijpers et al., 1998). Although electrophysiological studies may further illuminate the nature of this asymmetry (see Kooijman et al., submitted), it seems that Dutch infants need slightly more time to acquire their metrically based segmentation skills. This delay may be due to a difference in the contrast between strong and weak syllables. In English, unstressed syllables undergo more vowel reduction. This then increases the saliency of the strong syllables in the language, possibly providing a more salient metrical cue for English-learning infants. The less salient contrast in Dutch may require the infants to hear more of the language before they can discriminate between the different levels of stress. In addition, in Dutch, as well as in German and Spanish, but not in English, stress plays an important role in adult lexical recognition (Cutler & Pasveer, 2006). Considering the more complex stress structure of Dutch and the increased importance of stress for lexical processing, infants may need more time to fully learn all aspects of the Dutch stress pattern.

Although the Dutch language has a complex stress system, little is known about the development of the ability to segment weak-strong words from continuous speech. Only a few behavioral and electrophysiological studies have addressed word segmentation in Dutch (Kuijpers et al., 1998; Kooijman et al., 2005; submitted) and these have focused on strong-weak words. However, as pointed out above, many words in Dutch, as in English, have a weak-strong stress pattern. Here we present the first electrophysiological study of segmentation of weak-strong words in Dutch ten-month-olds and of the role of the strong syllable in this task. In a Familiarization and Test paradigm (Jusczyk & Aslin, 1995), we first familiarized ten-month-old infants with Dutch low-frequency weak-strong words, and then tested them on sentences containing the familiarized word or a strong-weak word with the same strong syllable. In addition, we presented control sentences with unfamiliar weak-strong and strong-weak words. The design of this experiment allowed us to study brain response to isolated weak-strong words, to the segmentation of weak-strong words, and to strong syllables in a different speech context.

Considering the ease with which the boundaries of words in isolation can be detected (silence is the clearest marker of a word boundary), we expect an ERP response to the isolated words similar to the response for strong-weak isolated words found by Kooijman et al. (2005). The response to the weak-strong words in the sentences, however, is less easy to predict. Dutch infants at ten months of age are quite proficient at word segmentation (Kooijman et al, 2005; Kuijpers et al., 1998), at least with strong-weak words. As described above, Jusczyk et al. (1999) have shown successful segmentation of weak-strong words in English-learning 10.5-month-olds. As also noted, however, Dutch infants seem to develop their word segmentation skills slightly later than English infants; English but not Dutch 7.5-month-olds succeed in segmenting strong-weak words (Jusczyk et al., 1999; Kuipers et al., 1998). Thus, it is not at all certain that Dutch ten-month-olds will be able to make use of multiple segmentation cues in the language. It is possible that in both the weak-strong and strong-weak words we will find an ERP response time-locked to the strong syllable only, if our ten-month-olds are at an earlier stage of development than Jusczyk et al.'s English learners. However, Jusczyk et al.'s results also show that the immediate context in which strong syllables occur is of crucial diagnostic importance. If infants are able to build a representation of the weak-strong words in isolation, this information may be used in different ways to recognize words in continuous speech. Differential processing of a repeated weak-strong word as compared to only a repeated strong syllable in sentence context may be indicated by differences in the ERP responses.

METHODS

Participants

Twenty Dutch ten-month-old infants from monolingual families participated in this study (mean age = 305 days; age range = 283-318 days; 8 female). Thirty-three additional infants were tested, but excluded from data analyses because, due to restlessness or sleepiness, not enough data could be collected. All infants were reported to have normal development and hearing, no major problems during pregnancy and birth, and no neurological or language problems in the immediate family. All infants were born in term. One infant had a left-handed half-brother; the other infants had no left-handedness in the immediate family. The parents signed an informed consent form and received 20 euro for participation.

Stimuli

Thirty-four bisyllabic words with main stress on the second syllable were selected from the CELEX Dutch lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). In addition, because not enough real words could be found which matched to a strong-weak word for pairing purposes, six further bisyllabic pseudowords with the same stress pattern were created (see *Appendix 1B* for a list of all forty word pairs and the corresponding sentences). All words were low in frequency. For each word and pseudoword a bisyllabic pair with stress on the first syllable was selected or a pseudoword with the same pattern was created. In the two words of a pair, the stressed syllable was the same. For example, *tij* in the pair *tijger* ('tiger') - *getij* ('tide'). For each target word, a set of two sentences was constructed (see Table 1 for an example). The word preceding the target word as well as the position of the target word in the sentences were matched across pairs. The target words in the sentences were never in first or final position. The stimuli were recorded digitally in a sound-attenuating booth by a female native Dutch speaker in a lively child-directed manner.

Design

The experiment consisted of forty experimental blocks. Each block consisted of a Familiarization phase comprising eight tokens of a weak-strong target word, and a Test phase consisting of four sentences. In half of the Test phases, two sentences contained the familiarized word the infant just heard during Familiarization, and two contained an unfamiliar weak-strong word from a different pair. In the other half, two sentences contained the paired strong-weak word (thus, the strong syllable was the same as in the weak-strong familiarization word), and two contained an unfamiliar strong-weak word from a different pair (see Table 2 for an example). The sentences in each Test phase were randomized. The mean length of the words was 1080 ms in the Familiarization phase, and 762 ms and 720 ms respectively for the weak-strong and strong-weak words in the sentences. Words spoken in isolation are naturally longer than words spoken in sentences, hence the difference in word length between the Familiarization phase and the Test phase. The small difference in length between the target words in the sentences is mostly due to final lengthening of the strong syllable in the weak-strong words. The sentences had a mean duration of 3190 ms.

Four lists were created, counterbalancing Test type (that is, in two of the four lists the weak-strong sentences are replaced by strong-weak sentences and vice versa); and Order of presentation (that is, two of the four lists were presented in reversed order). Each list was presented to five infants.

Table 1: An example of a word pair and its sentences.

Weak-strong word	*getij* ('tide')
Sentences	*Het wilde getij bedaart.* (The wild tide is calming down.) *Na het vrij rustige getij volgt storm.* (The fairly quiet tide is followed by a storm.)
Strong-weak word	*tijger* ('tiger')
Sentences	*De wilde tijger springt.* (The wild tiger is pouncing.) *Het lijkt een vrij rustige tijger te zijn.* (It seems to be a fairly quiet tiger.)

Table 2: Example of an experimental block with the conditions weak-strong familiar (example word *getij*) and weak-strong unfamiliar (example pseudoword *megeel*).

Familiarization phase	*getij* ('tide'; eight tokens)
Test phase	*Hij legt wat megeel in de la.* (He is putting some megeel in the drawer.) *Het wilde getij bedaart.* (The wild tide is calming down.) *Dat is megeel uit Egypte.* (That is megeel from Egypt.) *Na het vrij rustige getij volgt storm.* (The fairly quiet tide is followed by a storm.)

Procedure

The experiment took place in a sound-attenuating test booth. The infant sat in a child seat in front of a computer screen. The parent sat next to the child and listened to a masking CD through closed-ear headphones. The stimuli were presented via loud speakers placed in front of the infant. Screensavers, not synchronized with the stimuli, were shown to keep the infants interested and still. The child was allowed to play with a small silent toy. The experiment took 24.5 minutes. We presented as much of the experiment as possible, until the infant got too distracted to continue. Breaks were taken when necessary. All subjects heard at least 25 blocks.

EEG recordings and analyses

Infant-size Brain-Caps with 27 Ag/AgCl sintered ring electrodes were used. 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). 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. The EEG electrodes were referenced to the left mastoid online and re-referenced to linked mastoids offline. 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 DC high-impedance EEG amplifier using a band pass of 0.01-200 Hz and a sample rate of 500 Hz. Impedances of the reference and ground electrodes were kept below 5kΩ; impedances of the EEG and EOG electrodes were kept below 50kΩ. Seven electrodes (Fz, FCz, Cz, Pz, Oz PO7, PO8) were excluded from analysis due to excessive artifact. An offline filter of 0.1 – 30 Hz was used. The individual trials were aligned to the onset of the target words and to the onset of the second syllable of the weak-strong target words. Offline, the EEG signal was screened for artifact from 200 ms before to 800 ms after the acoustic onset of the target word and second syllable. Trials with artifacts were rejected.

Average waveforms were calculated for each condition for each subject. The mean number of trials in each subject average was 35.5 in the Familiarization phase and 14 in the Test phase. Time windows for the analyses were chosen based on visual inspection of the waveforms. The averaged ERP to the first two tokens in each Familiarization phase (isolated unfamiliar) were compared to the averaged ERP to the last two tokens in the each Familiarization phase (isolated familiar). In the Test phase, the ERPs to the repeated weak-strong words were compared to the unfamiliar weak-strong words; the ERPs to the repeated strong syllables (in the strong-weak words) were compared to the unfamiliar strong-weak words. In addition, we calculated average waveforms time-locked to the onset of the second syllable of the weak-strong words, because differences in the length of the preceding weak syllables may mask any ERP effects related to the strong, and thus more salient, syllable. Repeated measures analyses of variance statistics (ANOVA) were performed on the mean amplitudes in the selected time-windows with Familiarity (Familiar vs. Unfamiliar), Quadrant (4; Left Frontal, Right Frontal; Left Posterior; Right Posterior), and Electrode (5; Left Frontal: F7, F3, FT7, FC3, C3; Right Frontal: F8, F4, FT8, FC4, C4; Left Posterior: LT, LTP, CP3, LP, P3; Right Posterior: RT, RTP, CP4, RP, P4) as variables. For all tests, the Huynh-Feldt epsilon correction was used. The original degrees of freedom and adjusted p-values are reported. For significant effects, additional T-tests were performed on 50 ms windows with 40 ms overlap (e.g., 0-50 ms, 10-60 ms, etc.) to determine the onset of the effect. Significance on five consecutive 50 ms windows is considered evidence for onset.

RESULTS

Familiarization phase

The ERPs to the familiar words show a large negative-going deflection as compared to the unfamiliar words. The grand mean waveforms for the familiar and unfamiliar words start to diverge not later than 200 ms after word onset (see Figure 1). Analyses of the 200-500 ms time-window shows a main effect of Familiarity ($F(1,19)=15.1$, $p<.05$). (Also see *Supporting Table 3, Appendix 2B.)* Onset tests indicate this effect starts at 140 ms for electrode P4 and at 160 ms for electrodes FC3, FC4, and C4.

The results of the Familiarization phase show an effect of repetition similar to the effect found by Kooijman et al. (2005). They tested ten-month-olds on word segmentation of strong-weak words only. Each Familiarization phase consisted of ten tokens of the same bisyllabic strong-weak word. Both the onset and the duration of the effect we found for the weak-strong words are similar to the effect for the strong-weak words.

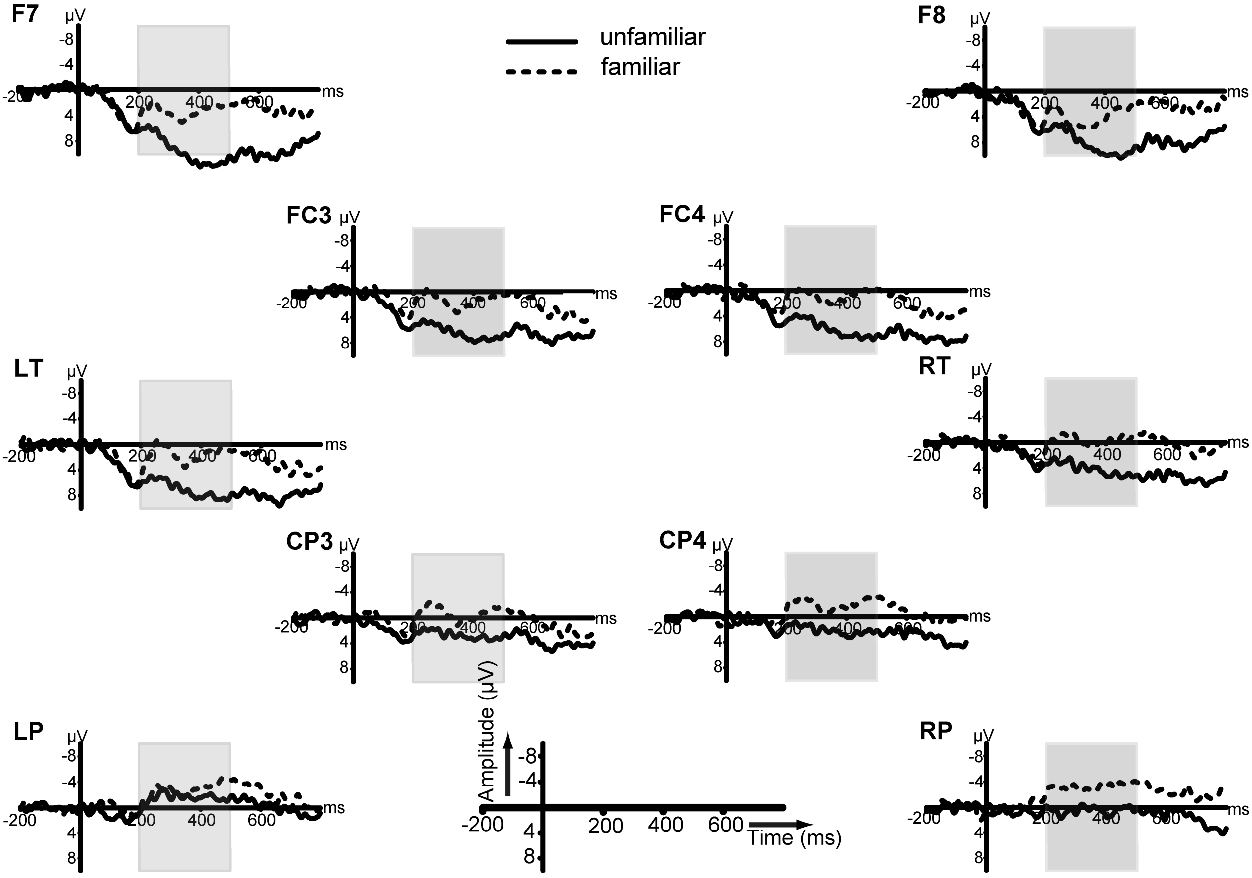

Figure 1: Familiarization phase. *The grand mean waveforms to the familiar and unfamiliar isolated weak-strong words on a subset of electrodes. The grey area indicates the time window selected for analyses (200 - 500 ms). Negativity is plotted upwards.*

Test phase

Weak-strong words

Inspection of the grand mean waveforms aligned to onset of the words shows a deviation between the familiar and unfamiliar weak-strong words in the sentences starting at about 600 ms (see Figure 2). Since this deviation is visible at the end of the chosen time window, we also calculated the mean waveforms time-locked to the onset of the second syllable (also see *EEG recordings and analyses* in the

Method section of this chapter*)*. The ERPs aligned to the onset of the second, strong, syllable of the words shows a larger effect than was seen in the waveforms aligned to word onset. Moreover, the effect is temporally less smeared out and has a clearer onset starting at about 370 ms (see Figure 3 and 5). This difference between the waveforms is confirmed by statistical analyses. Analyses in the 680-780 ms time window after word onset shows a marginally significant effect of Familiarity ($F(1,19)=3.41$, $p=.080$), whereas analyses in the 370-500 ms time window from onset of the second syllable show a significant effect of Familiarity ($F(1,19)=5.00$, $p<.05$). (Also see *Supporting Tables 4a and 4b, Appendix 2B.*) The onset analyses for the strong syllable show that this effect starts at 370 ms for electrodes F3, FT7, FC3, FT8, C3, RT and RTP.

Strong-weak words

The grand mean waveforms of the strong-weak words (aligned to the onset of the word) show a positive-going deflection to the familiar strong syllable in the 55-135 ms time window as compared to the unfamiliar strong syllable. This effect is smaller and has a different polarity than the effect to the strong syllable in the weak-strong words. In a later time window, from 300-500 ms, the familiar strong-weak words again show a positive-going deflection on several frontal electrodes (see Figure 4).

Analyses in the 55-135 ms time window revealed a significant interaction of Familiarity by Quadrant ($F(1,19)=3.07$, $p<.05$). Further analyses per Quadrant showed a main effect of Familiarity for the Right Frontal Quadrant ($F(1,19)=5.56$, $p<.05$). Onset tests show a significant onset starting at 40 ms for electrode F8.

For the 300-500 ms time window a significant interaction of Familiarity by Quadrant was found ($F(1,19)=3.59$, $p<.05$). Further analyses, however, did not show a main effect of Familiarity in any of the Quadrants. (Also see *Supporting Tables 5a and 5b, Appendix 2B.*) Additional analyses time-locked to the second syllable on the 200-350 ms time window did not reveal any significant effects ($F(1,19)<1$) either.

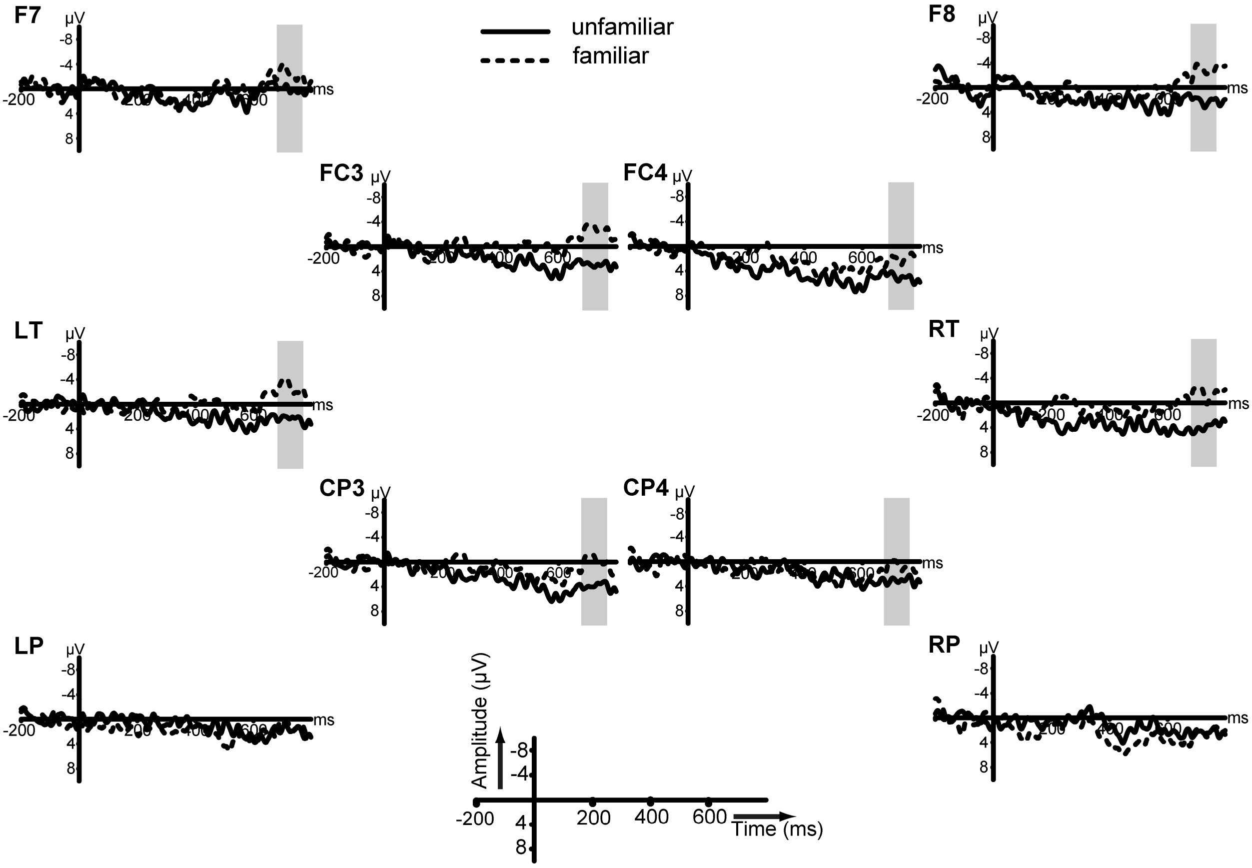

Figure 2: Test phase. *The grand mean waveforms aligned to the onset of the familiar and unfamiliar weak-strong target words in the sentences on a subset of electrodes. The grey area indicates the time window selected for analyses (680-780 ms). Negativity is plotted upwards; 0 ms is the onset of the target words.*

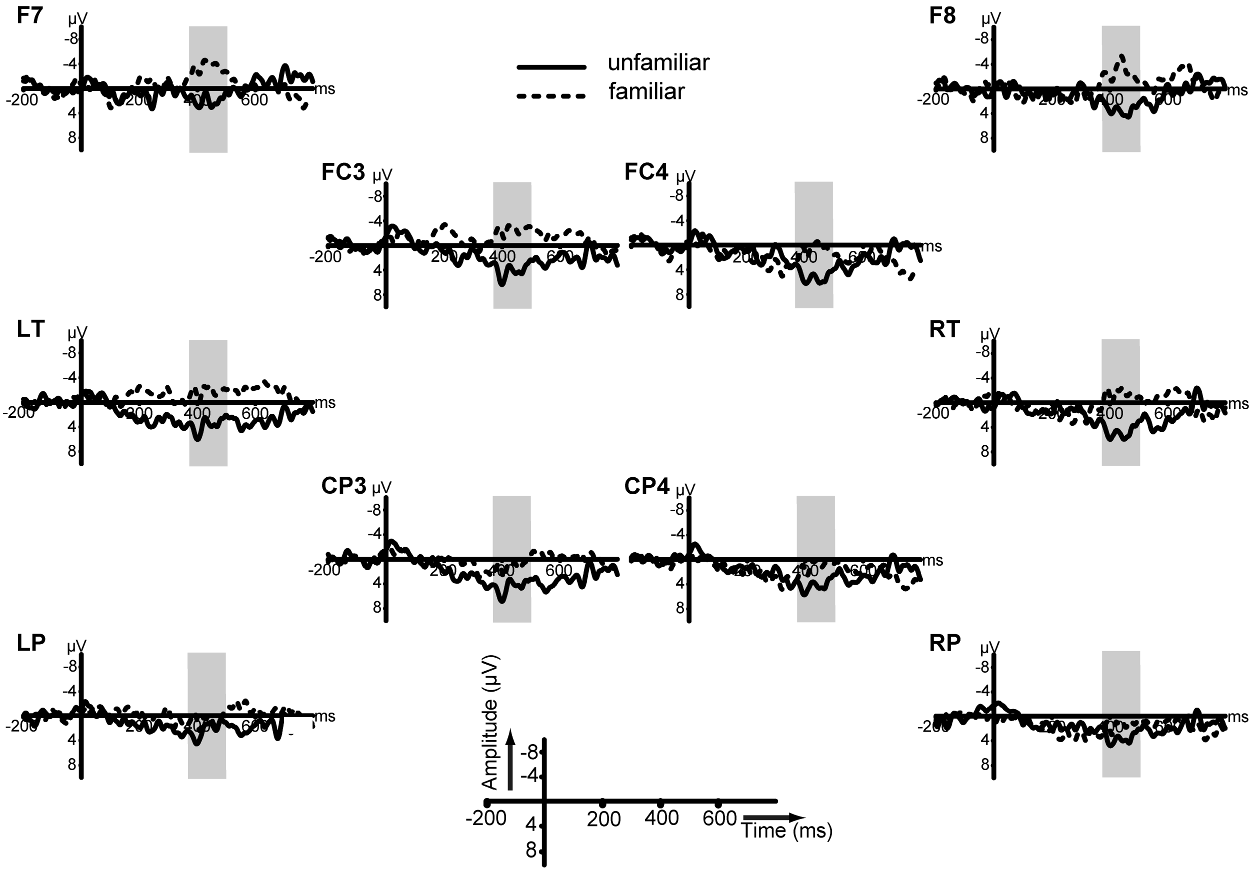

Figure 3: Test phase. *The grand mean waveforms aligned to the onset of the second, strong, syllable of the familiar and unfamiliar weak-strong target words in the sentences on a subset of electrodes. The grey area indicates the time window selected for analyses (370-500 ms). Negativity is plotted upwards; 0 ms is the onset of the second syllable of the target words.*

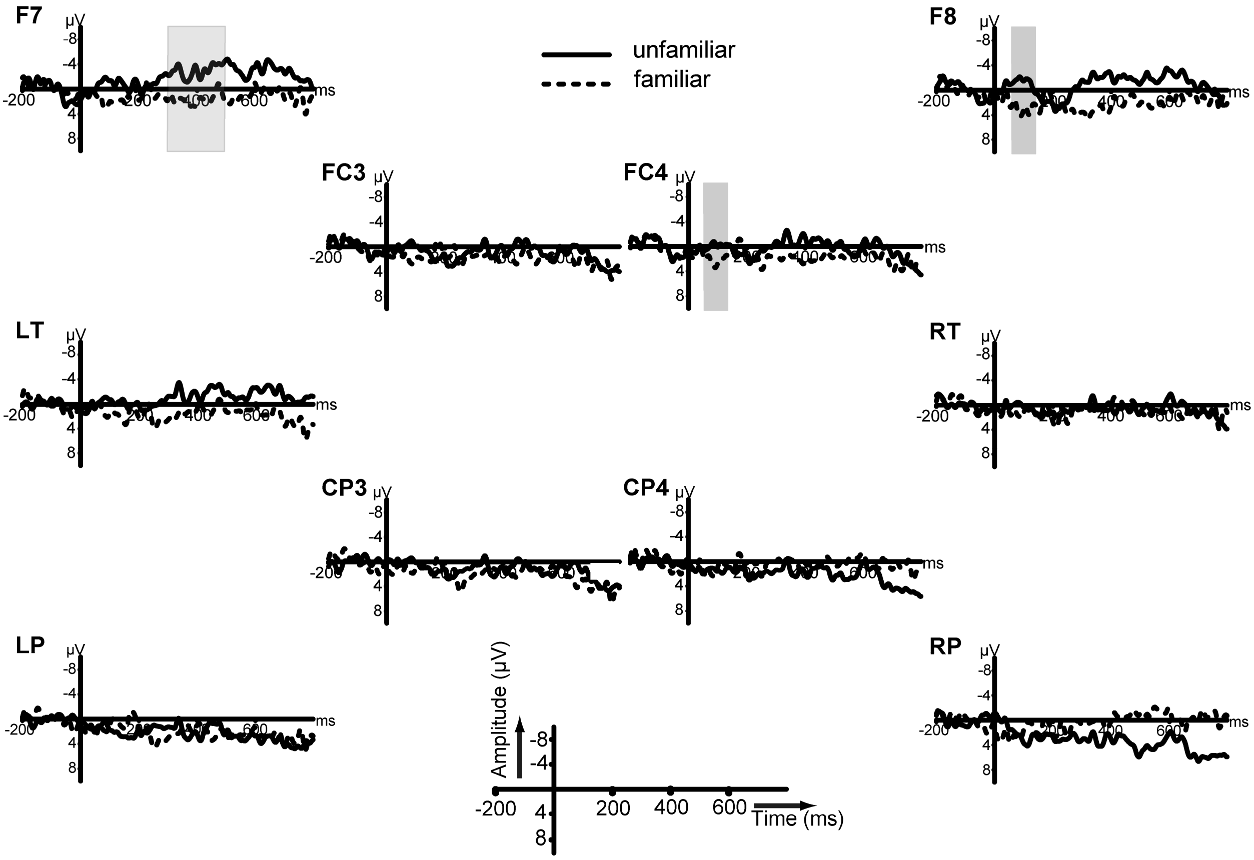

Figure 4: Test phase. *The grand mean waveforms to the familiar and unfamiliar strong syllables of the strong-weak target words in the sentences on a subset of electrodes. The grey area indicates the time window selected for analyses (F8, FC4: 55-135 ms; F7: 300-500 ms). Negativity is plotted upwards; 0 ms is the onset of the target words.*

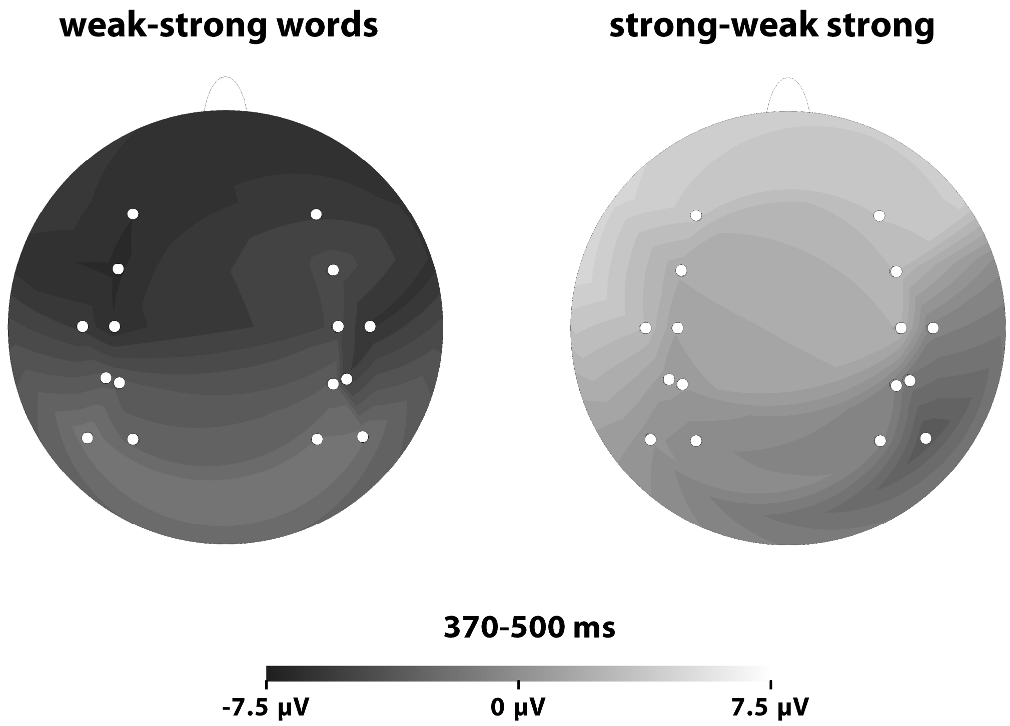

Figure 5: Isovoltage plots of the difference waves of the Test phase. *The difference waves are calculated by subtracting the ERP to the unfamiliar target words from the ERP to the familiar target words in the sentences. Figure 5 shows the isovoltage plots of the difference wave to the strong syllable of the weak-strong and strong-weak words in the 370-500 ms time window.*

GENERAL DISCUSSION

Dutch ten-month-old infants rely principally on strong syllables for word segmentation, even after familiarization with isolated weak-strong words. However, they are also sensitive to the context in which the strong syllables appear. Together, these results suggest that Dutch ten-month-olds are beginning to develop sensitivity to cues other than the most salient stress cue with which they first achieved segmentation.

In our experiment, we presented Dutch ten-month-olds with a Familiarization phase and a Test phase. In each Familiarization phase, the infants heard eight tokens of the same bisyllabic weak-strong word. The results to the Familiarization phase show a negative-going deflection to the (first) weak syllable of the familiar isolated words as compared to the unfamiliar isolated words.

Each Test phase of the experiment consisted of four sentences, either containing the familiar and an unfamiliar weak-strong word, or strong-weak words with the familiar or an unfamiliar strong syllable. The results show a familiarity response to the strong syllable of the weak-strong words from 370-500 ms, and to the strong syllable of the strong-weak words from 55-135 ms.

Thus, infants perceive the familiar strong syllables in continuous speech. Even when the infants heard the strong syllable only in a weak-strong word during Familiarization, they showed recognition of the strong syllable in a different context. This result suggests that Dutch ten-month-old infants still largely rely on the strong syllable for word segmentation.

This may seem surprising considering the brain response to the isolated weak-strong words. The results of the Familiarization phase of this experiment show that, as expected, listening to a repeated weak-strong word elicits a brain response in Dutch ten-month-olds similar to a repeated strong-weak word. The onset and direction of this response are comparable to the brain response for repeated strong-weak words found by Kooijman et al. (2005). In addition, it begins well before the end of the first, weak, syllable (onset at 160 ms, mean syllable length: 332 ms). Thus, the infants process the repeated isolated weak-strong words from word onset. These results suggest that infants form a memory trace of the whole weak-strong word and not just of the strong syllable of the word. This is not surprising, since silence is the clearest marker of a word boundary. Thus, the infants can easily perceive the first weak syllable, and do not need to rely on the strong syllable of the word only. Nevertheless, recognizing a word in isolation is not the same as locating its boundaries and extracting it from continuous speech. The strongest cue to a word boundary, silence, is only reliably present before words presented in isolation. In sentence context, infants have to rely on probabilistic word segmentation cues. Thus, recognition beginning from word onset when weak-strong words are presented in isolation does not automatically entail that word segmentation will then be initiated from the weak syllable when words are presented in sentence context.

However, the results of our study cannot simply be interpreted as indicating that the infants segment only strong syllables from continuous speech. The ERP responses to the strong-weak and weak-strong words differ considerably. First, the response to the strong-weak words starts earlier than to the weak-strong words. This early segmentation response to the strong-weak words may be partly due to between-word coarticulation cues. Infants at ten months of age are highly efficient at strong-weak word segmentation, and may already be able to make use of these early cues in combination with the metrical information in the sentence to find word onsets. Second, the ERP effect for the weak-strong words is larger and has an opposite polarity than the effect for the strong-weak words. Moreover, it is distributed over the whole head whereas the response to the strong-weak words is only present over the right frontal area. These differences indicate that partly different processes are going on in the different conditions. Note that in our experiment the familiar weak-strong words in the Test phase were full repetitions of the words presented in the Familiarization phase. This was different for the strong-weak words. Here, only the second syllable of the words that were presented in the Familiarization phase was repeated. If the infants only processed the strong syllable in the sentences, regardless of context, the effects to the different conditions would have looked similar. However, this is not the case. The larger effect to the weak-strong words indicates that it represents more than just a segmentation response. Repetition of a strong syllable in the same context may trigger not just a segmentation response but also a recognition response to more than just the strong syllable. Some of the information acquired during familiarization with the weak-strong words may be used in sentence context. However, this experiment was not designed to study different cues to word segmentation, other than the role of metrical stress. Further research is needed to disentangle possible different overlapping ERP responses to different cues in the speech signal.

Overall, these results show that the phonetic context adjacent to a strong syllable also matters to the infants. We suggest that Dutch ten-month-olds are beginning to combine word boundary cues. However, they are not yet fully in command of the segmentation procedures they need for dealing with weak-strong words, in that they do not show a rapid familiarity response to the initial weak syllable when it is surrounded by other syllables in continuous speech. In addition, they produce what is apparently a false-positive response to the same strong syllables in words in which they occur in initial position, although the difference in ERP signature in this latter case does suggest that the infants are sensitive to the difference in immediately adjacent phonetic context.

As we noted in the Introduction, the metrical structures of English and of Dutch are highly similar. But they are not quite similar enough for either adult or infant processing to run identical courses in the two languages. In English, extensive vowel reduction effectively amplifies the strong-weak differences, by grouping syllables into two more clearly differentiated categories. In Dutch, there is more gradation, and many syllables have full vowels but are unstressed. Pairs of cognate words make the asymmetry clear; *cobra* in English has a reduced final vowel, but *cobra* in Dutch has the unstressed but full final vowel /a/; and English *cigar* has a reduced first vowel, while Dutch *sigaar* has the unstressed but full first vowel /i/. There are significant consequences of this small asymmetry for both adult and infant listeners. Adult Dutch-speakers take suprasegmental cues to stress into account in word recognition, because it pays off for them to do so (Cutler & Pasveer, 2006), while adult English-speakers largely ignore stress in word recognition, apart from its use in segmentation (see Cutler, 2005, for a review).

In infancy, the cross-linguistic difference has the consequence that rates of development differ. Although the most effective initial segmentation cue is the same for each language – segmentation at the onsets of strong syllables – it is acted on earlier in the developmental trajectory of young English-learners than of young Dutch-learners. Successful segmentation of stressed monosyllabic words and of strong-weak words from continuous speech is observed in HPP experiments from 7.5 months in English-learning infants (Jusczyk & Aslin, 1995; Jusczyk et al., 1999), but at nine months and later in Dutch-learners (Kuijpers et al., 1998; Kooijman et al., 2005). The present results with weak-strong words are fully consistent with this pattern with strong-weak words. Jusczyk et al. showed that English-learning 10.5-month-olds resist false positive responses to, for instance, *tar* when they are presented with passages containing the word *guitar*. Our ten-month-olds are not as selective; their responses to words like *getij* in sentences were clearly dependent on the strong syllable, and were not launched by the initial weak syllable. (Note that the superior temporal sensitivity of the ERP technique allows us to observe this dependency in a single experiment, rather than needing to compare across experiments with bisyllables and monosyllables as would be the case with HPP.) Thus the Dutch ten-month-olds lag behind the course of English-learners' development. The Dutch infants' response, indeed, resembles the pattern that Jusczyk et al.'s HPP study recorded for 7.5-month-olds, who when familiarized with *guitar* failed to detect it in sentence context, but in two cases – monosyllabic familiarization and bisyllabic test, or the reverse – did produce false positive responses to *tar* when given *guitar*.

That the Dutch infants in our study appear somewhat further in development than Jusczyk et al.'s 7.5-month-olds, we argued, is evidenced by the fact that their false positive response to the initial syllables of words like *tijger* was different in kind to their response to the strong syllables in the *getij* words. We propose that Dutch-learning infants at an only slightly later age should show an ERP effect to the weak syllable in weak-strong words as well, analogous to the ability of Jusczyk et al.'s 10,5-month-olds to find a familiarized *guitar* in sentence context. English-learning 10.5-month-olds, of course, should already be able to show such an ERP response, although empirical evidence of this is not yet available.

The pattern of results across the various studies in English and Dutch thus shows a consistent lag between the two infant populations; the exploitation of prosodic cues to segmentation consistently occurs earlier in English. Similarly, weaker cues to phrasal juncture in Dutch than in English lead to a comparable lag in infants' use of these cues (Johnson & Seidl, submitted). By contrast, there is no obvious asymmetry in the accessibility of cues to segmental identity in the two languages, and infant sensitivity to mispronunciation of known words seems to be equivalent in English (Vihman, Nakai, DePaolis & Hallé, 2004) and Dutch (Swingley, 2005). The fine detail of language-specific phonological structure clearly exercises considerable influence on the course of development of particular language processing skills.

Finally, we note that to fully understand the development of word segmentation, we of course need to study adult word segmentation as well. So far, only a few studies have addressed this issue using electrophysiological techniques (Sanders & Neville, 2003; Snijders, Kooijman, Hagoort & Cutler, submitted). These studies focused on words with a strong initial syllable. We do not know yet what the adults' ERP response to weak syllables in continuous speech looks like, and whether they show a response time-locked to the weak syllable at all. This information is necessary to determine at which stage of development infants reach an adult-like level of word segmentation.

REFERENCES

Baayen, R.H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bergeson, T.R., & Trehub, S.E. (2006). Infants' perception of rhythmic patterns. *Music Perception 23*, 345-360.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in four-month-old infants from monolingual and bilingual environments. *Cognition 65*: 33-69.

Cutler, A. (2005). Lexical stress. In D.B. Pisoni & R.E. Remez (Eds.) *The Handbook of Speech Perception* (pp. 264-289). Oxford: Blackwell.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language 31*: 218-236.

Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language 2*, 133-142.

Cutler, A., & Pasveer, D., (2006). Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition. In *Proceedings of the Third International Conference on Speech Prosody*, Dresden, Germany.

Houston, D.M., Santelmann, L.M., & Jusczyk, P.W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes 19*: 97-136.

Johnson, E.K. (2005). English-learning infants' representations of word forms with iambic stress. *Infancy 7*: 99-109.

Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language 44*: 548-567.

Johnson, E.K. & Seidl, A.H. (submitted). A cross-linguistic perspective on infants' detection of junctures in speech. *Infancy*.

Jusczyk, P.W. (1999). How infants begin to extract words from fluent speech. *Trends in Cognitive Science 3*: 323-328.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology 29*, 1–23.

Jusczyk, P.W., Cutler, A., & Redanz, N.J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development 64*, 675-687.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology 39*: 159-207.

Jusczyk, P.W., Luce, P.A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language 33*: 630-645.

Jusczyk, P.W., Friederici, A.D., Wessels, J., Svenkerud, V.Y., & Jusczyk, A.M., 1993. Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*: 402-420.

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research 24*, 109-116.

Kooijman, V., Hagoort, P., & Cutler, A. (submitted). Word recognition at seven months: mismatching behavioral and electrophysiological evidence.

Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience 5*: 831-843.

Kuijpers, C., Coolen, R., Houston, D., & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. In: *Advances in Infancy Research*. Eds. C. Rovee-Collier, L. Lipsitt and H. Hayne. London: Ablex Publishing.

Mattys, S.L., & Jusczyk, P.W. (2001). Do infants segment words of recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance 27*: 644-655.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology 38*: 465-494.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language 39*: 21-46.

Morgan, J.L., and Saffran, J.R. 1995. Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development 66*: 911-936.

Myers, J., Jusczyk, P.W., Kemler Nelson, J., Charles-Luce, J., Woodward, A.L., & Hirsch-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language 23*: 1-30.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance 24*: 756-766.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication 41*, 233-243.

Newman, R., Bernstein Ratner, N., Jusczyk, A.M., Jusczyk, P.W., & Dow, K.A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology 42*, 643-655.

Norris, D., McQueen, J.M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition 21*: 1209-1228.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-olds. *Science 13*: 1926-28.

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research, 15*, 228-240.

Snijders, T., Kooijman,V., Hagoort, P., & Cutler, A., (submitted). Neurophysiological evidence of delayed segmentation in a foreign language.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science, 8,* 432-443.

Vihman, M., dePaolis, R., Nakai, S., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language, 50,* 336-353.

Van de Weijer, J. (1998). Language input for word discovery. PhD Dissertation, Radboud University Nijmegen. *MPI Series in Psycholinguistics 9*.

Weber, C., Hahne, A., Friedrich, M., & Friederichi, A.D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence. *Cognitive Brain Research 18*: 149-161.

Word recognition at seven months: mismatching behavioral and electrophysiological evidence

Chapter 4

This chapter is a slightly adjusted version of the paper Kooijman, V., Hagoort, P., & Cutler, A., submitted. Word recognition at seven months: mismatching behavioral and electrophysiological evidence.

Infants amass a substantial amount of information about the sound structure of their native language in the first year of life. They learn to extract words from continuous speech before they know the meaning of most words in their native language. In the absence of obvious absolute cues to word boundaries in continuous speech, they have to make use of the probabilistic cues that are available, such as metrical stress, and phonotactic and allophonic cues. Previous behavioral research showed word segmentation skills in Dutch nine-month-olds, but not in younger Dutch infants. In this paper we present electrophysiological evidence of word segmentation already at seven months of age. The infants show a differential brain response to familiarized two-syllable words in continuous speech as compared to unfamiliar words. In addition, we performed a behavioral head turn preference study using the same stimuli. The lack of a behavioral preference in this study indicates the brain response is a *precursor* of the corresponding behavioral response.

INTRODUCTION

Before infants speak their first words, they amass a substantial amount of information about the sound structure of the language they are exposed to (Bates, Thal, Finlay, & Clancy, 2002; Kuhl, 2004; Werker & Tees, 1999). Most of the language they are hearing, however, comes in the form of continuous speech (Van de Weijer, 1998). One of the earliest problems infants face, therefore, is how to extract possible words from this input. To adult listeners, who seemingly without effort hear the individual words in continuous speech, this segmentation task may appear unproblematic; but there are no consistent pauses between words in spoken sentences, and no absolute cues to word boundaries – certainly nothing as reliable as the cues printed text contains, in the form of spaces between words. Lexical cues are important for adult word segmentation (Norris, McQueen, Cutler & Butterfield, 1997), but are obviously not at the disposal of the beginning listener. Only probabilistic pre-lexical cues, such as the metrical stress pattern of a language, phonotactics (i.e., possible phoneme combinations in the language), and statistical regularities are available (Jusczyk, 1997; Mattys, Jusczyk, Luce, & Morgan, 1999; Saffran, 2001; Saffran, Aslin, & Newport, 1996). Infants in the first year of life have to discover these probabilistic cues in order to learn their native language.

Because most of the input consists of continuous speech, and not isolated words, infants cannot first learn words and then learn to deal with them in speech; rather they have to deal with continuous speech in order to learn potential words. The very fact that most infants produce words by the time they are about a year old shows that they do master this formidable task. By that time, infants have detected many of the pre-lexical segmentation cues in their native language, and are quite proficient at word segmentation. This is a very important step in language development. How important it is has recently been shown by Newman, Bernstein Ratner, Jusczyk, Jusczyk, and Dow (2006); they found that word segmentation skills before 12 months of age are directly predictive of vocabulary size at age two.

Much of what we know about word segmentation in infants comes from studies using the Head Turn Preference Procedure (HPP). Jusczyk and Aslin (1995) described an adjusted HPP method particularly suitable for studying infants' recognition of words in continuous speech. Their experiments showed that American English infants could segment monosyllabic words from speech at 7.5 months of age; six-month-olds, however, did not yet show this ability. Further HPP studies supported and extended these results, showing segmentation of bisyllabic (Jusczyk, Houston & Newsome, 1999) and trisyllabic (Houston, Santelmann, & Jusczyk, 2004) words with stress on the first syllable in 7.5-month-olds, and segmentation of weak-strong bisyllabic words in 10.5-month-olds (Jusczyk et al., 1999). The results of these studies suggest there is an important role for metrical cues in early word segmentation, at least in English; the majority of words in English have a strong-weak stress pattern, providing a possible cue to word segmentation. In other words, strong syllables indicate where word boundaries are likely to be in continuous speech. This is not the only cue in the language, but, at least in English and possibly in other stress-based languages, it is a very salient cue.

Studies similar to those of Jusczyk and his colleagues have also been performed in other trochaic stress-based languages, e.g., in Dutch (Kuijpers, Coolen, Houston, & Cutler, 1998) and in German (Höhle & Weissenborn, 2003, 2005; for an overview of segmentation studies in different languages, see Nazzi, Iakimova, Bertoncini, & de Schonen, in press). However, the results have not been exactly the same in these languages. In particular, one study directly analogous to that of Jusczyk et al. (1999) failed to show segmentation of strong-weak bisyllabic words in Dutch 7.5-month-olds (Kuijpers et al., 1998). Only at nine months of age did the Dutch infants show that they could segment such words from speech context. Kuijpers et al. observed that even though Dutch and English have very similar metrical stress patterns, English unstressed syllables undergo more vowel reduction. This effectively increases contrast between the syllables, whereby strong syllables become more salient in English than in Dutch. Such salience could facilitate segmentation of this word type for infants acquiring English, which in turn could explain the age difference in English versus Dutch infants' acquisition of initial segmentation skills.

Word segmentation in Dutch infants has also been examined with a different task; Kooijman, Hagoort and Cutler (2005) devised an Event Related Brain Potential (ERP) paradigm for this purpose. ERP is an online measure which has certain advantages over behavioral methods such as HPP: it has high time resolution, making it possible to study immediate effects as words are heard in continuous speech; and it does not require behavioral responses, making it a particularly useful technique to study cognitive processes in young infants. Kooijman et al. tested ten-month-olds. In their experiment, as in the segmentation experiments with HPP, infants were first familiarized with isolated words and then tested on sentences containing either these familiar words, or unfamiliar words. Their study revealed an ERP response to the familiar words in the sentences, well before the end of the words. Thus, ten-month-old infants show a differential brain response to familiar words as compared to unfamiliar words in continuous speech. These results are in line with the results of Kuijpers et al., and suggest that Dutch ten-month-olds only need roughly the first half of a strong-weak word to initiate word segmentation.

In this paper, we present two experiments addressing the word segmentation performance of younger Dutch infants. Experiment 1 uses the same ERP paradigm as Kooijman et al., to examine segmentation of strong-weak words by seven-month-old infants. This of course is the age group which, in the HPP experiments of Kuijpers et al., showed no evidence of word segmentation. Nevertheless, the online ERP measure, which does not require a behavioral response, may give a more direct reflection of Dutch infants' capacities at this age. It may reveal cognitive processes which have not matured enough to trigger a behavioral response. In Experiment 2 we then used the same materials as in the

ERP study in a HPP design, to determine whether Dutch seven-month-olds show a behavioral response to these particular stimuli.

METHODS: ERP STUDY

Participants

Twenty-eight seven-month-old infants from Dutch monolingual families participated (mean age = 218 days; age range = 194-232 days; 13 female). Twenty-two additional infants were tested, but excluded from data analyses because of fussiness or sleepiness. All infants were reported to have normal development and hearing, and no major problems during pregnancy or birth. All infants were full term, except for one, who was born 3.6 weeks premature. There were no neurological or language problems in the immediate families. Two infants had an older left-handed sister, and one had a mother who was forced right-handed; all others had no left-handedness in the immediate family. The parents signed a consent form and received 20 euro for participation.

Stimuli and design

The stimulus materials and design are the same as in Chapter 2 of this dissertation. Forty low-frequency bisyllabic nouns with main stress on the first syllable were selected from the CELEX Dutch lexical database (Baayen, Piepenbrock, and Van Rijn, 1993; for example: *zwaluw* ([zwa·lyw]; 'swallow'); or *viking* ([ví·kíŋ]; 'viking')). Twenty noun pairs were formed, and a set of four sentences was constructed for each noun. The word preceding the noun, as well as the position of the noun in the sentences were matched across pairs. For an example of a pair of nouns and its corresponding sentences, see Table 1, and *Appendix 1A*. The stimuli (ten tokens of each noun and four sentences per noun) were recorded by a Dutch female speaker in a lively child-directed manner. Recordings were made in a sound-attenuating booth onto digital audio tape. The recordings were sampled to disk at 16 kHz mono and edited using a speech waveform editor. The mean duration of the nouns was 710 ms for the isolated words (range: 373 – 1269 ms) and 721 ms for the target words in the sentences (range 224 – 1046 ms). The sentences had a mean duration of 4082 ms (range: 2697 – 5839 ms).

The experiment consisted of 20 experimental Familiarization and Test blocks (see Table 1 for an example), each containing 10 tokens of a target word (Familiarization), followed by eight randomized sentences (Test). Four of these sentences contained the word just familiarized; four contained the unfamiliar paired word. Four lists were created, counterbalancing Familiarization type (that is, in two lists half of the target words were used for Familiarization, the two other lists the other half of the target words were used for Familiarization) and Order of presentation (that is, two of the four lists were presented in reversed order). Each list was presented to seven infants.

Table 1: Example of an experimental trial in the ERP study

Familiarization	Ten tokens of <u>*zwaluw*</u> or <u>*viking*</u>
Test	*Een <u>zwaluw</u> vliegt vaak laag over het landschap.* (A swallow often flies low across the land.) *Die kleine <u>viking</u> is niet sterk maar slim.* (The little viking is clever but not strong.) *Een <u>viking</u> gaat op reis naar verre landen.* (A viking travels to far countries.) *Ik zie een andere <u>zwaluw</u> in de wei.* (I see another swallow in the meadow.) *Pieter zag die <u>viking</u> uit het noorden.* (Pieter saw the viking from the north.) *Dat is die andere <u>viking</u> met veel vijanden.* (That is the other viking with many enemies.) *'s Ochtends is die <u>zwaluw</u> altijd erg actief.* (The swallow is always very active in the mornings.) *De kleine <u>zwaluw</u> kan heel goed vliegjes vangen.* (The little swallow is good at catching flies.)

Procedure

The experiment took place in a sound-attenuating test booth. The infant sat in a child seat in front of a computer screen and listened to the stimuli presented via three loudspeakers placed in front of the child. A screensaver, not synchronized with the stimuli, was shown to keep the infants interested. In addition, the infants were allowed to play with a small silent toy. The parent sat next to the child and listened to a masking CD through closed-ear headphones. The experimenter controlled the stimuli using the NESU (Nijmegen Experiment Setup) stimulus presentation program. We presented as many of the Familiarization and Test blocks as possible, until the infant got too distracted to continue. The experiment took about 32 minutes; the mean length of the blocks was 1.6 minutes, with 2.5 seconds of silence between the isolated words and 4.2 seconds of silence between the sentences. Breaks were taken when necessary. All subjects heard at least eight blocks (mean=13, range: 8 - 20); each block consisted of a Familiarization phase (ten tokens) and a Test phase (eight sentences).

EEG recordings

Infant-size Brain-Caps with 27 Ag/AgCl sintered ring electrodes were used to measure the Electroencephalogram (EEG). 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). 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. The left mastoid was used as the online reference for all electrodes. The EEG electrodes were referenced to the left mastoid online and 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 DC EEG amplifier using a band pass of 0.1-30 Hz and a sample rate of 200 Hz. Two occipital electrodes (PO7, PO8), as well as the midline electrodes (Fz, FCz, Cz, Pz, Oz), were excluded from analysis due to excessive artifact. Impedances were around 10 kΩ at the remaining 20 electrodes (see Figure 1 for the final electrode arrangement). Offline, individual trials were aligned 200 ms before the acoustic onset of the target words, and screened for artifact from -200 to 800 ms. Trials with artifacts were rejected (70% and 75% respectively for the isolated words and the target words in the sentences; these percentages are based on the entire experiment). Mean waveforms were calculated for each condition for each subject in the -200-800 ms window. The mean number of trials in each subject mean waveform after artifact rejection was 11.4 for the Familiarization phase and 19.6 for the Test phase. From subject mean waveforms, grand mean waveforms per condition were calculated. Time windows for statistical analyses were chosen based on visual inspection of the data. The number of trials used in each grand mean waveform was 324 and 293 for the unfamiliar and familiar isolated words respectively, and 550 and 554 for the unfamiliar and familiar target words in the sentences respectively. To get rid of excess slow wave activity common in young infants which may obscure possible ERP effects (see Weber, Hahne, Friedrich & Friederici, 2004), we filtered the EEG signal offline to 1-30 Hz prior to further analyses. Repeated measures analyses of variance statistics were performed for these time windows with Familiarity (Familiar vs. Unfamiliar), Quadrant (4; Left Frontal, Right Frontal; Left Posterior; Right Posterior), and Electrode (5; Left Frontal: F7, F3, FT7, FC3, C3; Right Frontal: F8, F4, FT8, FC4, C4; Left Posterior: LT, LTP, CP3, LP, P3; Right Posterior: RT, RTP, CP4, RP, P4) as independent variables. For all tests, the Huynh-Feldt epsilon correction was used. The original degrees of freedom as well as the adjusted p-values are reported. The onsets of the effects were tested by performing t-tests on the difference waveforms on bins of 50 ms with an overlap of 40 ms (i.e., 0-50, 10-60 etc), whereby significance from zero ($p<.05$) on five consecutive bins is considered evidence for onset.

RESULTS AND DISCUSSION: ERP

Isolated words

The isolated words offer the opportunity to establish the presence of sensitivity to repetition. We averaged the EEG to token 1 and 2 of the familiarization phase, representing the ERP response to the most unfamiliar isolated words, and the EEG to token 9 and 10, at which point the infants had already heard eight tokens of the same word, representing the ERP response to the most familiar of the isolated words. We then compared these two averages; a difference between them is indicative of the infant recognizing the repetition.

The ERPs to the familiar versus the unfamiliar words differ in the 200-500 ms time window, mostly over the frontal electrodes (see Figure 1). In addition, two early peaks are more negative to the familiar words than to the unfamiliar words: from 40-120 ms (N1) over a subset of electrodes; from 220-320 ms over almost all electrodes. We analyzed the mean amplitudes in these time windows. The N1 did not show significant differences ($p>.05$).

In the 220-320 ms window, a main effect of Familiarity was found ($F_{1,27}=4.64$, $p=.04$). There was no significant interaction of Familiarity by Quadrant ($p>.05$). (Also see *Supporting Table 3, Appendix 2C*). Thus, the ERP effect of Familiarity is equally distributed over the head.

In the 200-500 ms window, we found a significant interaction of Familiarity x Quadrant ($F_{1,27}=2.7$, $p=.05$). Analyses per Quadrant revealed a main effect of Familiarity over the Left Frontal Quadrant ($F_{1,27}=6.15$, $p<.05$). No significant effects ($p<0.5$) were found for the Right Frontal and Posterior Quadrants (also see *Supporting Table 3, Appendix 2C*). Thus, the broad negative ERP effect to the familiar isolated words is strongest over the left frontal area.

Onset analyses (see *Methods*) revealed an onset starting at 220 ms for the electrodes F7 and FT7.

These ERP results show a brain response to the repetition of tokens of the same word starting at 220 ms. This Familiarity response is similar to, but later than that found in the study of Kooijman et al. (2005), in which ten-month-olds showed a Familiarity response starting at 160 ms. Just like the ten-month-olds, however, the present seven-month-old listeners can recognize repetition of the same form in isolation, a prerequisite for being able to detect repetition of the same form in a speech context.

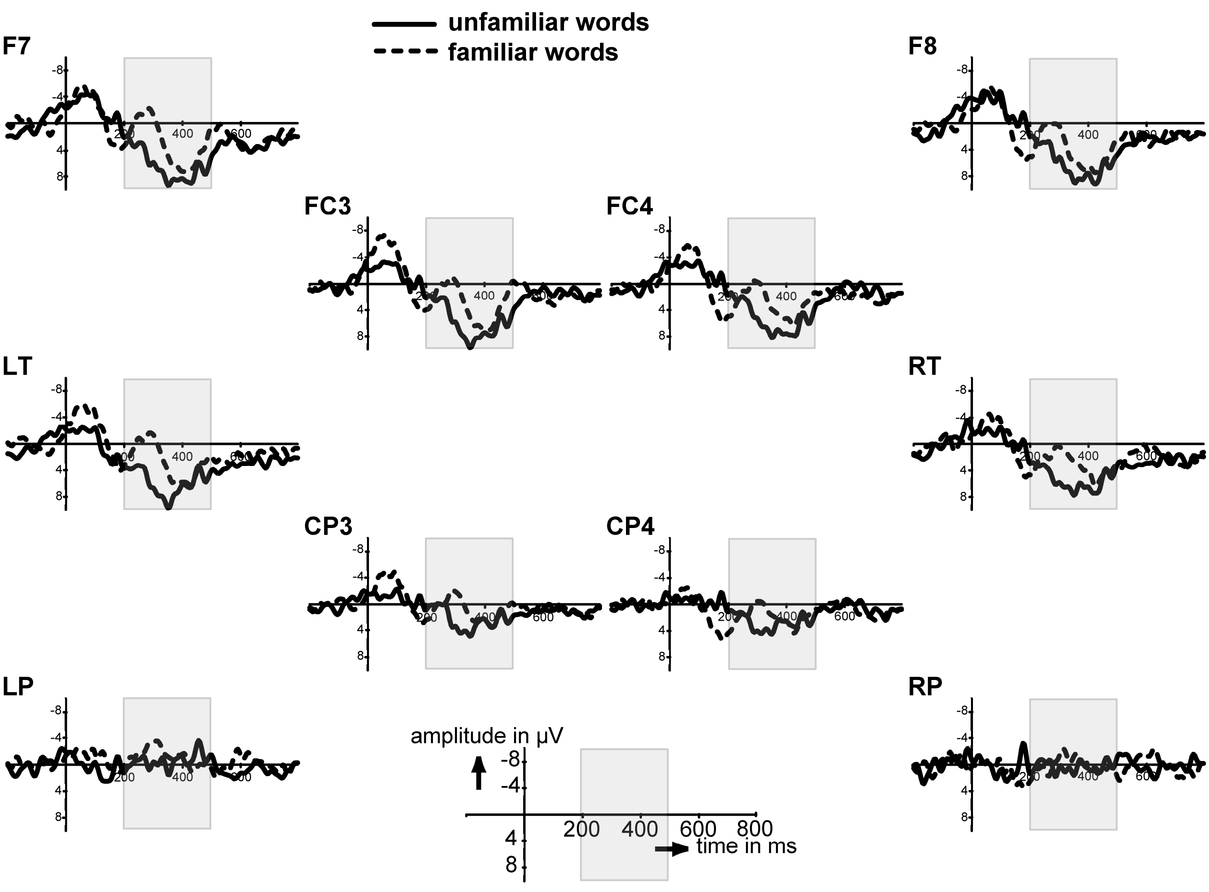

Figure 1: Familiarization phase. *The grand mean waveforms to the unfamiliar (word position 1/2) and familiar (word position 9/10) isolated words on a subset of electrodes; negativity is plotted upwards. The grey area indicates the time window of 200 to 500 ms.*

Sentences

We calculated the ERPs to the familiar and unfamiliar words in the sentences. The grand mean waveforms deviate over the frontal areas from 350 to 450 ms, and over the left posterior area starting at about 430 ms with an opposite polarity as to the frontal effect (see Figure 2). We performed statistical analyses over the mean amplitudes in the time windows 350-450 ms (see Figure 2A) and 430-530 ms (see Figure 2B).

A significant interaction of Familiarity x Quadrant ($F_{1,27}$=4.05, p<.05) was observed for the 350-450 ms window. Analyses per Quadrant showed a marginally significant effect of Familiarity over the Right Frontal Quadrant ($F_{1,27}$=2.7, p=.065). This result suggests that the effect is present over a more restricted area of the brain within the Right Frontal Quadrant. Therefore, we performed further analyses over a subset of four electrodes (F4, F8, FC4, and FT8) in that quadrant, which revealed a significant main effect of Familiarity ($F_{1,27}$=4.3, p<.05; also see *Supporting Table 4, Appendix 2C*).). There were no significant effects (p<.05) in equivalent analyses for the Left Frontal or Left or Right Posterior Quadrants. Thus, the early effect of Familiarity is strongest over the right frontal brain area. Onset tests (see *Methods*) revealed a significant effect (p<.05) at 300 ms for electrode FT8.

Visual inspection of the grand mean waveforms in the 430-530 ms window shows that the effect in this window is restricted to electrodes over the left hemisphere at the posterior sites LTP, CP3, and P3. To test this local effect, we only included this subset of left posterior electrodes in the analysis. We found a significant effect of Familiarity ($F_{1,27}$=4.2, p<.05) over these three electrodes.

The results of the ERP study indicate that at seven months of age, Dutch infants can detect words previously heard in isolation when they re-occur in continuous speech. These results differ from previous behavioral results (Kuijpers et al., 1998) that showed no evidence of word segmentation at 7.5 months of age. This difference in results may reflect a difference in sensitivity between the paradigms used. However, it is also the case that different stimuli were used in the two studies. Thus, it is possible that the present stimuli were spoken more slowly or with more pronounced intonation than those of Kuijpers et al., making word segmentation easier for our participants. To investigate the possibility of behavioral segmentation responses to our stimuli, we designed a HPP study, using the same materials as in the ERP study.

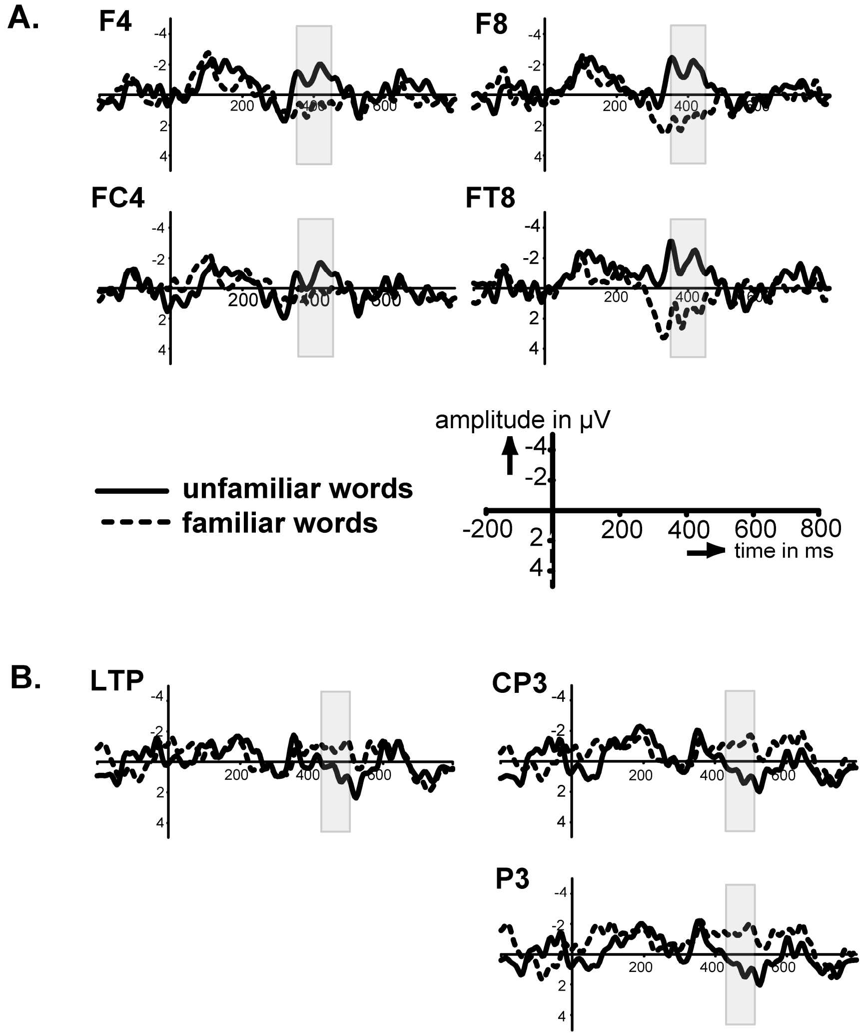

Figure 2: Test phase. *The grand mean waveforms to the familiar and unfamiliar target words in the sentences on a subset of electrodes; negativity is plotted upwards; filter 1-30 Hz. Figure 2A: a subset of Right Frontal electrodes. The grey area indicates time window of 350-450 ms. Figure 2B: a subset of Left Posterior electrodes. The grey area indicates the time window of 430-530 ms.*

METHODS: HPP STUDY

Participants

Twenty-eight seven-month-old infants (mean age = 212 days; age range= 198-228 days; 12 female) from Dutch monolingual families participated. The infants had normal development and hearing, and no major problems during pregnancy or birth. One infant was born 15 days premature; the others were full term. One infant had a dyslexic father, and one a father and brother who were dyslexic; the others had no language problems in the immediate family. The parents received five euro or a present of a toy for their participation.

Stimuli and design

Ten pairs of bisyllabic nouns and the corresponding sentences were selected from the EEG stimuli. We used a slightly adapted version of the HPP study of Jusczyk and Aslin (1995), with ten consecutive blocks (instead of one block as is normally the case in HPP studies). Each block consisted of ten tokens of the same word (Familiarization), followed by four trials of four sentences each (Test): two trials with the familiarized word in each of the four sentences, and two trials with its unfamiliar pair (see Table 2).

The increased number of Familiarization and Test blocks, as well as the design of Familiarization, closely resembled the ERP study which required a high number of experimental trials to reach an adequate signal-to-noise ratio. The Test blocks were closer to the original HPP study, in that there were four consecutive trials in different conditions instead of randomized sentences. Four versions of the experiment were created as in the ERP study. Each version was presented to seven infants.

Table 2: Example of an experimental trial in the adjusted HPP study.

Familiarization	Ten tokens of <u>*zwaluw*</u> or <u>*viking*</u>
Test	*Een <u>viking</u> gaat op reis naar verre landen.* *Die kleine <u>viking</u> is niet sterk maar slim.* *Dat is die andere <u>viking</u> met veel vijanden.* *Pieter zag die <u>viking</u> uit het noorden.*
	Een <u>zwaluw</u> vliegt vaak laag over het landschap. *De kleine <u>zwaluw</u> kan heel goed vliegjes vangen.* *Ik zie een andere <u>zwaluw</u> in de wei.* *'s Ochtends is die <u>zwaluw</u> altijd erg actief.*
	Een <u>viking</u> gaat op reis naar verre landen. *Die kleine <u>viking</u> is niet sterk maar slim.* *Dat is die andere <u>viking</u> met veel vijanden.* *Pieter zag die <u>viking</u> uit het noorden.*
	Een <u>zwaluw</u> vliegt vaak laag over het landschap. *De kleine <u>zwaluw</u> kan heel goed vliegjes vangen.* *Ik zie een andere <u>zwaluw</u> in de wei.* *'s Ochtends is die <u>zwaluw</u> altijd erg actief.*

Procedure

The experiment took place in a three-sided booth. Infants sat on a caregiver's lap facing the center panel of the booth. The test booth had a red light attached at eye level to the center panel and a blue light attached to each side panel. A camera was mounted behind the center panel under the red light, with its lens through a hole in the panel. The experimenter observed the infant on a monitor connected to the camera. A computer and a response box were situated behind the center panel for stimulus presentation. The experimenter used the response box to start and stop the stimuli, and relay information on the direction and duration of the head turns to the computer. The infant could not see the experimenter behind the center panel. During the experiment, the experimenter and the caregiver listened to masking music over closed-ear headphones.

The stimuli were presented from loudspeakers mounted behind the light on each side panel. During Familiarization, the sidelights flashed contingent upon the infants' looking behavior. The lights were not linked to the presentation of the Familiarization stimuli. Following each Familiarization, the corresponding Test trials were presented. The trials were alternated, and played equally often from the two speakers while the light on the respective side was flashing. Looking time of the infant in the direction of the stimulus was measured. If the infant looked away for more then two consecutive seconds, the trial was ended, and the next trial or block started. If the infant continued to look in the direction of the stimulus, the trial was played to the end. The experiment was continued for as long as the infant was interested. Each infant heard at least three blocks. However, the results showed considerably reduced looking times after the first block. Therefore, we report the results of the first block here; this is directly analogous to a standard HPP design.

There were four versions of this first block. Half of the infants were familiarized with the Dutch word *zwaluw* ([zwa·lyw]; 'swallow'; n=7) or *viking* ([ví·kIŋ]; 'viking'; n=7), and tested on sentences containing these words (n=14). The other half were familiarized with the Dutch word *pudding* ([pΨdIŋ]; 'pudding'; n=7) or *sauna* ([sAu·na]; 'sauna'; n=7), and tested on sentences containing these words (n=14). For each group, the trials with sentences containing the familiarized words were then compared with trials with sentences containing the other word. The looking times to each trial were summed and inspected. Looking times below 2020 ms (i.e., mean onset + mean length of the target words) were excluded from analysis. The average looking times (looking times to blocks in each condition, familiarized versus unfamiliar, summed and divided by two in each case) were subjected to repeated measures analysis of variance.

RESULTS AND DISCUSSION: HPP

Thirteen of 28 infants showed a longer looking time to the test trials in the familiar condition, and thirteen to the unfamiliar condition; two showed no difference (see Figure 3). We performed repeated measures analyses on the looking times, with Familiarity as a dependent variable, and Version as a between-subjects factor. No significant differences were found for Familiarity ($F_{1,24}$=1.46, p=.24) or for the interaction (Familiarity x Version: $F_{1,24}$=.79, p=.51). (Also see *Supporting Table 3, Appendix 2C.*)

Although the same materials were used as in the ERP study, the HPP study does not show a preference response (see Figure 4), confirming previous HPP indications that Dutch seven-month-olds may not be able yet to segment words from continuous speech. In the next section, we will discuss the results of both studies and a possible conclusion that could only come from evidence of converging methods.

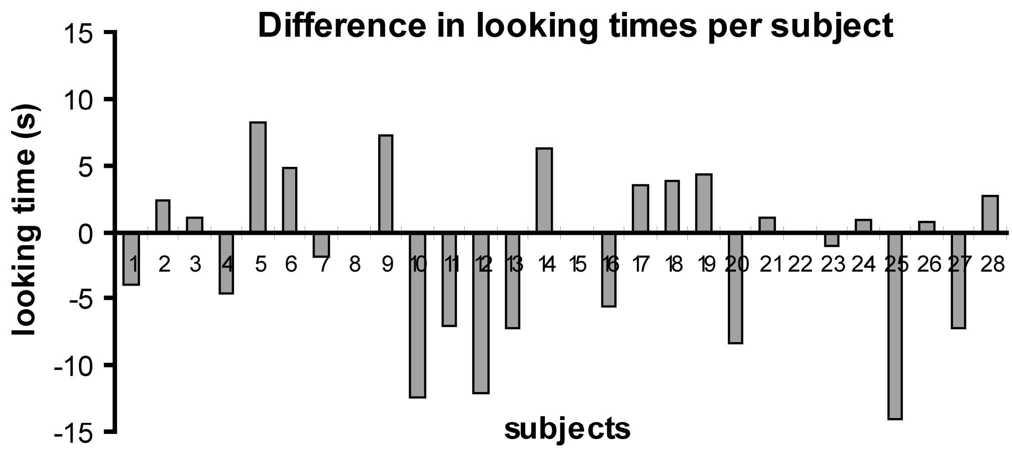

Figure 3: HPP experiment. *The difference in looking times per subject in block 1. The difference is calculated by subtracting the looking times to the sentences with the unfamiliar target words from the looking times to the sentences with the familiar target words. Positive values indicate a longer looking time for the familiar target words. Negative values indicate a longer looking time for the unfamiliar target words.*

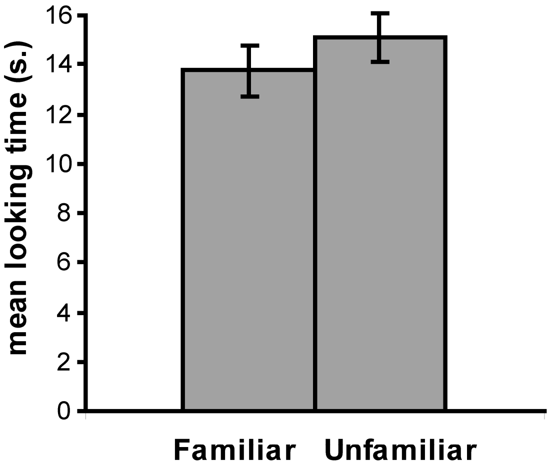

Figure 4: HPP experiment. *Mean looking times across subjects in block 1 for sentences with familiar versus unfamiliar target words.*

GENERAL DISCUSSION

Our ERP results provide evidence of word segmentation from continuous speech in Dutch seven-months-olds. At 350 ms after word onset, the infants show a differential brain response to the familiar words as compared to the unfamiliar words in the sentences. This right frontal effect starts roughly half way through the target words (recall that the mean length of these words was 721 ms), suggesting that seven-month-olds can initiate segmentation rapidly, e.g., from the first strong syllable. Following this early effect, the results show a small left centroparietal effect starting around 480 ms. Although it is as yet unclear whether these two effects reflect differential contributions to the segmentation process, it *is* clear that, in contrast to the results of previous behavioral studies, they indicate the presence of segmentation skills.

The polarity and distribution of the early Familiarity effect in our seven-month-olds and in Kooijman et al.'s (2005) ten-month-olds differ, however. This suggests that at least partly different processes are going on in these age groups. We suggest that one possible factor underlying the different ERP effects across age groups could be change in cognitive abilities. Research in different areas of development has shown that between eight and ten months of age, infants learn to combine different sources of information (Jusczyk, 1999; Morgan & Saffran, 1995; Werker et al., 1999). At seven months of age, infants may be only able to use one source of segmentation information; for English and Dutch infants this might be, for example, word stress. Around ten months of age, however, infants can combine different cues such as metrical stress and phonotactic and allophonic patterns (Jusczyk, Hohne, & Bauman, 1999; Juszcyk, Luce, & Charles-Luce, 1994). This allows for more efficient segmentation. In consequence, a stronger ERP effect of Familiarity appears in the ten-month-olds, as they can more efficiently extract discrete units from continuous speech. On this account, seven-month-olds' segmentation would rely on fewer cues, resulting in a less efficient and focused response.

Other factors may also cause differences in ERP polarity and distribution between these age groups. Physical factors, such as the closing of the fontanels (Flemmings, Wang, Caprihan, Eiselt, Haueisen, & Okada, 2005), as well as neural development, such as dendritic growth (Uylings, 2006), continue well after birth. Further research is needed to find out how these different factors (i.e., cognitive, neural and physical development) affect the polarity and distribution of the ERP signal early in life.

The latter left lateral effect found in the seven-month-olds is more similar in direction and distribution to the effect found in ten-month-olds (Kooijman et al, 2005). We suggest that this small effect is an early appearance of the stronger effect in the ten-month-olds. Some seven-month-olds may also already be more proficient at word segmentation. Note that a high variability in development in this age range was also found by Rivera-Gaxiola, Silva-Pereyra and Kuhl (2005) for phonetic discrimination.

Our HPP results show no evidence of word segmentation in Dutch learning seven-month-olds. The looking times are the same for the sentences containing familiar and unfamiliar words, even though these were the very same stimuli which did produce a significant ERP difference, the subjects were from the same population, and the experiments were kept as similar as possible. Though this null result parallels the outcome of previous behavioral studies of word segmentation in Dutch infants (Kuijpers et al., 1998), it thus seems at variance with our own ERP results. It might be suggested that an explanation for this asymmetry could be that seven-month-old Dutch infants are as yet unable to produce HPP responses. That is to say, the infants can segment words from speech, as shown by the ERP results, but the HPP is too hard a task for them to demonstrate this behavior. This suggestion can be discarded given that other HPP studies have shown significant effects with Dutch infants as young as six months (Johnson & Seidl, 2005).

We suggest instead that, even though Dutch infants at this age are sensitive to pattern recurrence in speech input, this sensitivity is not strong enough yet to prompt the corresponding behavioral response. We argued above that infants are first sensitive to segmentation cues such as those provided by the metrical stress pattern of their native language (see also Weber et al., 2004, for converging infant evidence from German, which is phonologically similar to Dutch and English). Our own ERP results suggest that they can apply this sensitivity even in continuous speech. The translation from this sensitivity to control of behavior, however, requires some, as yet undetermined, aspect of further development.

Note that a similar apparent conflict between brain and behavioral responses has been observed by McLaughlin, Osterhout and Kim (2004). They found N400 modulation in adult second language learners after only 14 hours of classroom training, even though the learners' performance on a word discrimination task was not above chance. They suggested that ERPs might be more sensitive to continuous change in knowledge than some behavioral methods. A similar conclusion may be drawn from our own results: Although the corresponding behavior in the HPP study is not yet present, the seven-month-olds' ERP response to familiar words in continuous speech shows that word segmentation skills are on the way. Thus, we suggest that the ability to segment words from speech shows up in our ERP results as a *precursor* of the corresponding behavioral response to word segmentation.

Further research using ERP is necessary to describe infants' brain response to continuous speech more closely. ERP is a valuable tool for studying language development in general, and word segmentation in particular, as it requires no behavioral response and has high sensitivity to time-course information. At the same time, behavioral methods, which have already provided us with the majority of what we know about language development, remain invaluable to give us a fuller picture of how brain and behavior relate, as the HPP has done in this study.

REFERENCES

Bates, E., Thal, D., Finlay, B., and Clancy, B. (2003). Early language development and its neural correlates. In: S. J. Segalowitz, & I. Rapin (Eds.) *Handbook of Neuropsychology:Vol. 8 part II Child Neuropsychology* (2nd ed., pp. 525-592). Amsterdam: Elsevier.

Baayen, R.H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexicaldatabase* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Flemming, L., Wang, Y., Caprihan, A., Eiselt, M., Haueisen, J. & Okada, Y. (2005). Evaluation of the distortion of EEG signals caused by a hole in the skull mimicking the fontanel in the skull of human neonates. *Clinical Neurophysiology 116*, 1141-1152.

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science 6*, 122-127.

Höhle, B., & Weissenborn, J. (2005). Word segmentation in German-learning infants. *Proceedings of the international workshop: Early word segmentation: a cross-linguistic approach taking advantage of Europe's linguistic diversity*, Paris, France.

Houston, D.M., Santelmann, L.M., & Jusczyk, P.W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes, 19*, 97-136.

Johnson, E.K., & Seidl, A. (2005). Cross-linguistic differences in 6-month-olds' clause segmentation strategies. *Proceedings of the ISCA workshop: Plasticity in Speech Perception*. London, UK.

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences 3*, 323-328.

Jusczyk, P.W. & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology 29*, 1-25.

Jusczyk, P. W., Hohne, E.A., & Bauman A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics 61*, 1465-76.

Jusczyk, P.W., Houston, D.M. & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology 39*, 159-207.

Juszcyk, P.W., Luce, P.A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language 33*, 630-645.

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research 24*, 109-116.

Kuijpers, C., Coolen, R., Houston, D. & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. In: C. Rovee-Collier, L. Lipsitt and H. Hayne (Eds.). *Advances in Infancy Research* (pp. 205-220). London: Ablex Publishing Corporation.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* 5, 831-43.

Mattys, S. L., Jusczyk, P.W., Luce, P.A., & Morgan, J.L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology* 38, 465-494.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience* 7, 703-704.

Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development 66,* 911-936.

Nazzi, T., Iakimova, G., Bertoncini, J., & de Schonen, S. (in press).

Behavioral and electrophysiological exploration of early word segmentation in French: distinguishing the syllabic and lexical levels. In: A.D. Friederici & G. Thierry (Eds.). *Trends in Language Acquisition Research Vol. 5*. Amsterdam: Benjamin Publishers.

Newman, R., Bernstein Ratner, N., Jusczyk, A.M., Jusczyk, P.W., & Dow, K.A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Developmental Psychology 42*, 643-655.

Norris, D., McQueen, J.M., Cutler, A., & Butterfield, S. 1997. The Possible-Word Constraint in the segmentation of continuous speech. *Cognitive Psychology 34*: 191-243.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P.K. (2005). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental Science* 8: 162-172.

Saffran, J. R. (2001). Words in a sea of sounds: the output of infant statistical learning. *Cognition 81*, 149-169.

Saffran, J. R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science 274*, 1926-1928.

Uylings, H.B.M (2006). Development of the human cortex and the concept of 'critical' or 'sensitive' periods. *Language Learning 56*(s1), 59-90.

Van de Weijer, J., (1999). *Language input for word discovery.* Unpublished doctoral dissertation, MPI Series in Psycholinguistics Vol. 9, MPI, Nijmegen, The Netherlands.

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence. *Cognitive Brain Research 18*, 149-161.

Werker, J. F., & Tees, R.C. (1999). Influences on infant speech processing: toward a new synthesis. *Annual Review of Psychology 50*, 509-535.

Neurophysiological evidence of delayed segmentation in a foreign language

Chapter 5

This chapter is a slightly adjusted version of the paper Snijders, T., Kooijman, V., Hagoort, P., & Cutler, A., in press. Neurophysiological evidence of delayed segmentation in a foreign language. Brain Research.

Previous studies have shown that segmentation skills are language-specific, making it difficult to segment continuous speech in an unfamiliar language into its component words. Here we present the first study capturing the delay in segmentation and recognition in the foreign listener using ERPs. We compare the ability of Dutch adults and of English adults without knowledge of Dutch (‘foreign listeners’) to segment familiarized words from continuous Dutch speech. We use the known effect of repetition on the event-related potential (ERP) as an index of recognition of words in continuous speech. Our results show that word repetitions in isolation are recognized with equivalent facility by native and foreign listeners, but word repetitions in continuous speech are not. First, words familiarized in isolation are recognized faster by native than by foreign listeners when they are repeated in continuous speech. Second, when words that have previously been heard only in a continuous-speech context re-occur in continuous speech, the repetition is detected by native listeners, but is not detected by foreign listeners. A preceding speech context facilitates word recognition for native listeners, but delays or even inhibits word recognition for foreign listeners. We propose that the apparent difference in segmentation rate

between native and foreign listeners is grounded in the difference in language-specific skills available to the listeners.

INTRODUCTION

"Parlez plus lentement, s'il vous plaît", "Bitte, sprechen Sie langsamer","Hable más despacio, por favor": Such utterances are the common resource of listeners attempting to understand an unfamiliar language: "Please, speak more slowly". Continuous speech contains no silences between words analogous to the spaces in written text. But while the continuity of spoken utterances is hardly noticeable in the native language, so that we effortlessly interpret each utterance as a sequence of individual words, the process of resolving continuous speech into words is markedly harder in a foreign language. This may explain why speech in foreign languages often seems unnervingly fast (Pfitzinger and Tamashima, 2006).

The difficulty of segmenting foreign speech lies in part in the language-specificity of the procedures by which listeners segment speech into words (Cutler et al., 1983; Cutler et al., 1986; Cutler et al., 1989; Dumay et al., 2002; Kolinksy et al., 1995; Otake et al., 1993; Suomi et al., 1997). Native listeners efficiently combine the prosodic, phonotactic and lexical cues and statistical regularities in the language to extract words from speech. The non-native listener, however, may be unable to call on the strategies of this kind which native listeners find to be effective. In part, segmenting foreign speech is also difficult because native segmentation procedures may be applied to other languages irrespective of whether they are appropriate (Cutler, 2000-2001; Cutler et al., 1986; Cutler and Otake, 1994; Otake et al., 1993; Vroomen et al., 1998). And finally, the native listener's ability to exploit syntactic and discourse information for rapid disambiguation will far outstrip that of the non-native listener. All these factors might combine to slow the segmentation process for non-native listeners.

However, it is currently unknown how great the difference in segmentation ability is. In this study we addressed this issue via on-line electrophysiological measures. We tested twelve native Dutch-speaking adults, and twelve native English-speaking adults without knowledge of Dutch, on segmentation of Dutch. We will refer to the latter group as the foreign listeners. Foreign listeners cannot call on any of the language-specific sources of knowledge that the Dutch listeners command. They have, in effect, as little working knowledge of the language as infant listeners, who are known to develop the ability to extract word forms from continuous speech before they start to learn word meanings (Jusczyk, 1999; Jusczyk and Aslin, 1995; Kooijman et al., 2005). Note, however, that the foreign listeners can in this case use partly similar segmentation procedures, as Dutch resembles English in the metrical structure called upon in segmentation (Cutler and Butterfield, 1992; Cutler and Norris, 1988; Vroomen et al., 1996). Our comparison thus allows us to focus on the effect of knowledge of the language on the ability to extract word forms from continuous speech.

Our study exploited the known effect of repetition on event-related brain potentials (ERPs): the ERP to a later presentation of a word is typically more positive than the ERP to the first presentation of the same word (Rugg, 1985; Rugg and Doyle, 1994; Rugg et al., 1995). Participants received twenty trials, each made up of two phases: Familiarization plus Test. In each Familiarization phase, ten tokens of a low-frequency Dutch word were presented in isolation. The words were all bisyllabic words with stress on the first syllable (e.g. *hommel*, 'bumble bee'). This type of word form is extremely common in both English (Cutler and Carter, 1987) and Dutch (Vroomen et al., 1996), and with one exception, the words conformed to English constraints on permissible syllable structures. In Familiarization, comparison of ERPs to the first versus the second token tests for a repetition effect for isolated words.

In each following Test phase, participants heard eight short sentences, of which half contained the familiarized word, and half a matched novel word. Table

1 shows an example of an experimental Test block (*hommel*, 'bumble bee', with its matched control *mammoet*, 'mammoth'; see *Appendix 1A* for all the materials). Familiarized status of the word tokens was counterbalanced across participants. The recognition of familiarized words in continuous speech was assessed by comparing the difference between ERPs to the first occurrence of the familiarized and the first occurrence of the unfamiliarized word in the sentences. In addition, ERPs to the first and the second presentation of the unfamiliarized word in continuous speech were compared to examine repetition effects to words that had previously been heard only in a sentence context (novel word repetition within Test).

Table 1: Example of one experimental block. Materials were in Dutch

Familiarization phase:
hommel hommel hommel hommel hommel
hommel hommel hommel hommel hommel

Test phase:
1. *Die kleine mammoet*[1] *zwemt in de rivier.*
   (That little mammoth swims in the river.)
2. *De hommel*[2] *vliegt van bloem naar bloem.*
   (The bumblebee flies from flower to flower.)
3. *Er is een oude mammoet*[3] *in het museum.*
   (There is an old mammoth in the museum.)
4. *De mammoet is al lang geleden uitgestorven.*
   (The mammoth became extinct long ago.)
5. *Vaak kan een hommel erg hard zoemen.*
   (Often a bumblebee can buzz very hard.)
6. *Het is een oude hommel met gele strepen.*
   (It is an old bumblebee with yellow stripes.)
7. *Daar is een mammoet met veel vriendjes.*
   (Over there is a mammoth with many friends.)
8. *Een kleine hommel zit op het gordijn.*
   (A little bumblebee is sitting on the curtain.)

[1]First unfamiliarized control word
[2]First familiarized word
[3]Second unfamiliarized word

To control for possible differences in memory load between the two groups, we conducted a second experiment, differing from Experiment 1 only in that pauses of 100 ms were inserted between the words in the sentences. This manipulation reduced the speech segmentation load, while the working memory load was kept constant. Since the familiarization phase was identical in the two experiments, we collapsed the familiarization results of both experiments.

EXPERIMENTAL PROCEDURE

Experiment 1

Subjects

Native language participants were twelve right-handed native speakers of Dutch (7 female, mean age 22, range 18-28 years). Foreign language participants were twelve right-handed native speakers of English (7 female, mean age 22, range 19-27). Six of them spoke British English and six American English. At the time of testing these subjects had been in the Netherlands for on average 2.4 months (range 1-7 months). They were unable to speak or understand Dutch. The answers of the English subjects on a Dutch lexical decision task did not differ from chance ($t = 1.97$, $p = 0.074$, mean = 54% correct, SD = 7.5 %). They could translate on average not more than 3.3 of 72 English monosyllabic words (e.g., *rope, sweep*) into Dutch. None of the participants had any neurological impairment or had experienced any neurological trauma according to their responses on a questionnaire. All subjects gave written informed consent.

Materials

Forty low frequency, two-syllable nouns with a strong/weak stress pattern were selected from the CELEX Dutch lexical database. These were arbitrarily formed into twenty pairs. For each of the forty nouns, a set of four sentences containing the noun was constructed. The position of the critical noun in the sentence and the word preceding it were matched within pairs. The sentences were short and contained, prior to the occurrence of each critical word, no semantic information that could have enabled native listeners to predict the word. Words and sentences were recorded in a sound-attenuating booth onto digital audiotape by a female native Dutch speaker, sampled at 16 kHz mono to disk, and edited using a speech waveform editor. The ten tokens of each word were acoustically highly variable. The mean duration of the words was 710 ms (range: 365-1270 ms) in isolation, 720 ms (range: 225-1045 ms) in sentence context. The mean sentence duration was 4080 ms (range: 2700-5840 ms).

Experiment 2

Subjects

Native language participants were twelve right-handed native speakers of Dutch (7 female, mean age 21, range 18-25 years). Foreign language participants were twelve right-handed native speakers of English (8 female, mean age 23, range 19-27). Five of them spoke British English and seven American English. At the time of testing these subjects had been in the Netherlands for on average 2.3 months (range 1 week - 8 months). They were unable to speak or understand Dutch. The answers of the English subjects on a Dutch lexical decision task did not differ from chance ($t = 0.28$, $p = .785$, mean = 50.5 % correct, SD = 6.5 %). They could translate on average not more than 1.7 of the same 72 English monosyllabic words into Dutch. None of the participants had any neurological impairment or had experienced any neurological trauma according to their responses on a questionnaire. All subjects gave written informed consent.

Materials

The materials were identical to the materials of Experiment 1. However, in Experiment 2 the words that made up the sentences were recorded separately, and the original sentences were reconstructed by concatenating these words, with 100 ms silence between adjacent words.

The mean duration of the words was 710 ms (range: 365-1270 ms) in isolation, 800 ms (range: 450-1190 ms) in sentence context. The mean sentence duration was 6030 ms (range: 4200-8170 ms).

Procedure

The procedure in both experiments was the same. The experimental trials were presented in 20 experimental blocks, each consisting of 10 different tokens of the same word (familiarization stimuli) followed by eight randomized sentences (test stimuli). Four of these contained the familiarized word (repetition condition), the other four contained the paired word, which had not been familiarized (non-repetition condition). Table 1 shows an example of an experimental block. Each block lasted approximately 1.6 minutes. There were short breaks between the blocks. In the Familiarization phase the different tokens of the same noun were separated by a silent interval of 2500 ms. In the Test phase, there was a silent interval between sentences of 4200 ms. Four versions of the experiment were constructed, such that the same nouns (and sentences) appeared in both the familiarized and the unfamiliarized conditions, and the presentation order of the blocks was counterbalanced. Thus in the Table 1 example, for half the listeners *hommel* was familiarized and *mammoet* was not, while for the other half *mammoet* was familiarized and *hommel* was not. EEG was measured during both the Familiarization and the Test phase. During EEG measurement the subjects were seated in a comfortable chair in front of a computer screen, in a dimly illuminated sound-attenuating booth. The subjects listened to the stimuli via a loudspeaker set, placed approximately 1.5 m in front of them. On the computer screen, a fixation asterisk was presented during the auditory presentation of the words and the sentences. The subjects were asked to avoid eye- and other movements during stimulus presentation. No additional task demands were imposed.

EEG recordings

EEG was measured using a BrainCap with 27 sintered Ag/AgCl electrodes. Twenty-one electrodes were placed according to the 10% standard system of the American Electroencephalographic Society (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). Another six electrodes were placed bilaterally on non-standard intermediate positions. A temporal pair (LT and RT) was placed 33% of the interaural distance lateral to Cz, while a temporo-parietal pair (LTP and RTP) was placed 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) was placed 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. Electro-oculogram (EOG) was recorded from electrodes above and below the eye, and at the outer canthi of the eyes. EEG and EOG data were recorded with a BrainAmp AC EEG amplifier using a high cut-off of 30 Hz and a time constant of 10 s. Impedances were typically kept below 3 kΩ for the EEG recordings and below 5 kΩ for the EOG recordings. The EEG and EOG signals were digitized online with a sample frequency of 200 Hz.

Data analyses

Individual trials were time-locked to the acoustic onset of the critical words. All trials were screened for eye movements, electrode drifting, amplifier blocking, and EMG artifacts, in a time window ranging from 200 ms before onset of the critical word to 1200 ms after the critical word. Trials containing artifacts were rejected. For the remaining trials a baseline correction was applied, in which the waveforms were normalized relative to a 100 ms stimulus-preceding epoch. Subsequently, averaged waveforms were computed. Statistical analyses of the repetition effects consisted of repeated measures analyses of variance (ANOVAs), using mean amplitude values for the 400-900 ms latency window computed for each subject, condition, and electrode site. To investigate the topographical distribution of the ERP-effects, different subsets of electrodes were grouped together (Anterior Left (AL): F7, F3, FT7, FC3, LT; Anterior Right (AR): F4, F8, FC4, FT8, RT; Posterior Left (PL): LTP, CP3, LP, P3, PO7; Posterior Right (PR): CP4, RTP, P4, RP, PO8). Omnibus 2 x 2 x 4 repeated measures ANOVAs on mean ERP amplitude (in μV) for the 400-900 ms time window were carried out first, with Group (native language, foreign language) as between-subject factor and Repetition (repetition/no-repetition) and Quadrant (AL, AR, PL, PR) as within-subject factors. When significant Repetition by Group interactions were found, separate ANOVAs were performed for the different groups. Where interactions between Repetition and Quadrant were significant, ANOVAs on the 4 quadrants were carried out separately. For the Familiarization phase, ERPs of Experiment 1 and the Experiment 2 were analyzed together, with Experiment as an additional between-subjects factor, as this phase was identical for both experiments (confirmed by absent Repetition by Experiment interactions, see *supporting Table 2, Appendix 2D*). For the Test phase ERPs of the two experiments were analyzed separately. For evaluation of effects with more than one degree of freedom in the numerator, the Greenhouse-Geisser correction was used. The original degrees of freedom and adjusted *p* values are reported.

To establish onset and duration of the repetition effect, cluster randomization analyses were performed using Fieldtrip, an open source toolbox for EEG and MEG analysis developed at the F.C. Donders Centre for Cognitive Neuroimaging (http://www.ru.nl/fcdonders/fieldtrip). The cluster randomization method that Fieldtrip uses is an improved version of the method described in Maris (2004) (Maris, 2004; Takashima et al., 2006). This test effectively controls the Type-1 error rate in a situation involving multiple comparisons (i.e., 27 electrodes x 240 time points). Briefly, the method works as follows: In a first step, all (electrode, time point-) pairs are identified for which the t-statistics for the difference between conditions (e.g. familiarized vs. unfamiliarized) exceed some prior threshold. The selected (electrode, time point-) pairs are then grouped into a number of clusters in such a way that, within every cluster, the (electrode, time point-) pairs form a set that is connected spatially and/or temporally. Each cluster is assigned a cluster-level test statistic whose value equals the sum of the (electrode, time point-) specific test statistics. Thus, the cluster-level test statistic depends on both the extent of the cluster and the size of the (electrode, time-) specific t-statistics that belong to this cluster. The Type-I error rate for the complete spatiotemporal data matrix is controlled by evaluating the cluster-level test statistic under the randomization null distribution of the maximum cluster-level test statistic. This randomization null distribution is obtained by randomizing the order of the data (e.g. familiarized and unfamiliarized trials) within every participant. By creating a reference distribution from 1000 random draws, the p-value may be estimated by the proportion from this randomization null distribution in which the maximum cluster-level test statistic exceeds the observed cluster-level test statistic (this proportion is called a Monte Carlo p-value in the statistics literature). With this number of 1000 random draws, our Monte Carlo p-value is an accurate estimate of the true p-value. In brief, the cluster randomization p-value denotes the chance that such a large summed cluster-level statistic will be observed when there is actually no effect. In this way significant clusters extending both over time and over electrodes can be identified, providing a measure both of the timing and of the distribution of the effect.

First, cluster randomization tests were performed to check for Repetition by Group interactions, comparing the size of the repetition effect for the native and the foreign listeners. Where interactions between Repetition and Group were significant, cluster randomization analyses to test the Repetition effect were carried out for native and foreign listeners separately. When no significant Repetition by Group interaction was found, both groups were analyzed together.

For illustrative purposes only, the grand mean ERPs were smoothed off-line using a 5-Hz low pass filter.

RESULTS AND DISCUSSION

Familiarization phase

The results showed a similar ERP response in the Familiarization phase for both the native and the foreign listeners: a positive repetition effect with a central-posterior distribution (see Figure 1). In the 400-900 ms time-window there was a significant effect of repetition ($F(1,44)=74.42$, $p=.000$) that was larger over posterior sites ($F(3,132)=33.30$, $p=.000$), and did not differ for the two groups ($F(1,44)=1.22$, $p=.276$, *Supporting Table 2a, Appendix 2D*). Onset analysis showed that the Repetition effect started at 240 ms (see *Supporting Table 2b, Appendix 2D*). Thus, both participant groups were equally able to recognize that the string of isolated tokens (e.g., *hommel, hommel, hommel...*) consisted of repetitions of the same word type. Prior knowledge of the language in which the words are spoken makes no difference to the nature of this response. This is consistent with previous research observing the same ERP repetition effect not only with words but also with pseudowords (Rugg et al., 1995), suggesting that no lexical knowledge is required for the appearance of this effect.

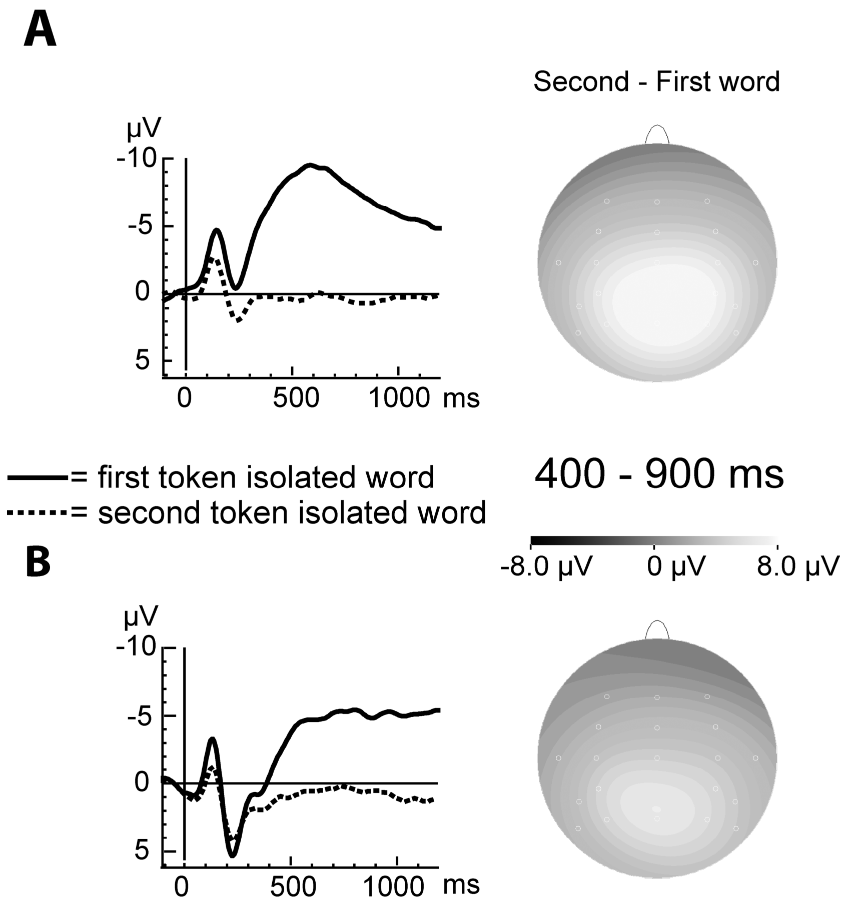

Figure 1: Familiarization phase. *Repetition Effect in the Familiarization phase for native (A) and foreign listeners (B). Left: Event-related potential (ERP) to the first and the second token of the word at a representative electrode site (Cz). Negativity is plotted upwards. Right: Topographic isovoltage maps of the single word repetition effect in the 400 – 900 ms latency range.*

Experiment 1: Test Phase

In the Test Phase of Experiment 1, however, ERP responses for the native and foreign listeners differed. Figure 2 shows ERPs to the first familiarized word and the first and second presentation of the unfamiliarized word in the sentences, for each group separately.

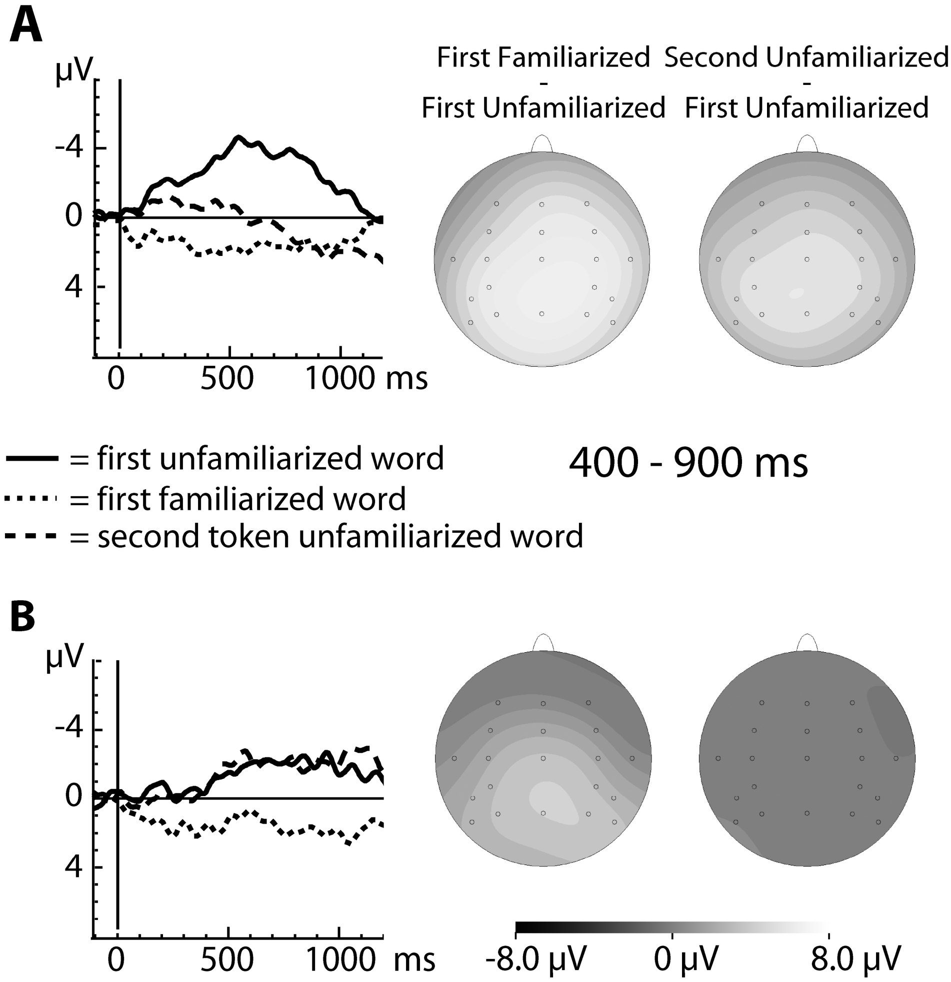

Figure 2: Test phase, Experiment 1. *Repetition Effect in the Test phase for native (A) and foreign listeners (B). Left: Event-related potential (ERP) to the first familiarized word, and the first and the second occurrence of the unfamiliarized control word in the sentences at a representative electrode site (Pz). Negativity is plotted upwards. Middle and Right: Topographic isovoltage maps of the different repetition effects in the 400 – 900 ms latency range. Middle: recognition of familiarized words in continuous speech: familiarized - unfamiliarized. Right: repetition effects within continuous speech: second unfamiliarized - first unfamiliarized.*

It can be seen that *native* listeners (Figure 2a) showed a repetition effect both to the familiarized words encountered in sentences ($F(1,11)=23.95$, $p=.000$), and to novel word repetition within Test ($F(1,11)=11.05$, $p=.007$, *Supporting Table 3, Appendix 2D*). The *foreign* listeners (Figure 2b) detected the occurrence of the familiarized word in the sentences ($F(1,11)=18.98$, $p=.001$), although their ERP repetition effect was reduced and substantially delayed (starting at 515 ms) compared to that of the native listeners (which started at 115 ms, *Supporting Table 4, Appendix 2D*). However, foreign listeners showed no repetition effect at all ($F<1$) in the comparison of first and second presentation of the *un*familiarized word in continuous speech (novel word repetition within Test). Detecting word forms in continuous speech was thus exceptionally difficult for foreign listeners.

We observed that the native listeners achieved segmentation from the preceding context and launched the recognition response rapidly - well within the time-span of the word's delivery. The mean duration of the two-syllable words in the sentences was 721 ms, and yet for familiarized words the native listeners initiated the segmentation and recognition process already at 115 ms. Thus, the process began well before the end of the first (stressed) syllable. Since the effect in continuous speech started 125 ms earlier than when the same words were presented in isolation, contextual cues may have helped native listeners to detect the repetitions in continuous speech. These contextual cues can presumably be similarly exploited whenever adult listeners segment their native language. As the example in Table 1 illustrates, our materials in general afforded no semantic or lexical cues which would have enabled the native listeners to anticipate the upcoming word. Thus, the cues in question could involve word-to-word coarticulation, syntactic structure, and rhythmic and prosodic predictability. The consequence of the native listeners' efficient use of this information is that as soon as the initial sounds of the familiarized word were heard, segmentation could take place, allowing word recognition to be initiated.

Consistent with this suggestion of rapid response to word-initial sounds is a finding of Sanders and Neville, who measured ERPs evoked in native listeners by different syllables in continuous speech; their experiments revealed a larger early sensory component (N100) for word-initial than for word-medial sounds (Sanders and Neville, 2003a; Sanders and Neville, 2003b; Sanders et al., 2002). In our experiment the familiarized words were strongly primed and expected to occur in the sentences, facilitating both the segmentation and the recognition process. Note that for novel word repetition within Test the repetition effect started only at 420 ms for native listeners; here the continuous speech context did not facilitate segmentation and recognition.

The pattern that we observed for foreign listeners in the Test phase differed from the native pattern. For familiarized words repeated in continuous speech a repetition effect occurred, but only from 515 ms. Novel word repetition in continuous speech, however, was not detected by these listeners. Thus with sufficient familiarization, foreign listeners could segment and recognize words in the sentence (although the repetition effect was delayed compared to that of native listeners); but without familiarization, segmentation and recognition did not occur at all. In other words, a preceding speech context helps native listeners but appears to hinder foreign listeners.

Word segmentation versus memory load

The results from Experiment 1 suggest that foreign listeners have difficulties recognizing words in continuous speech. Is this due to the segmentation difficulties they encounter, or to a larger working memory load (compared to the native listeners)? Native listeners can chunk the different words of the meaningful sentences into larger units, whereas foreign listeners can only store the unknown word forms individually. To investigate the possibility that our results in Experiment 1 were due to differences between the two groups in memory load rather than in segmentation capacities, we conducted a second experiment. In this experiment we used the same materials as in Experiment 1. However, in Experiment 2 the sentences for the Sentence Test phase were constructed from words spoken in isolation and concatenated, with 100 ms pauses between words. In this way segmentation is rendered unnecessary, while working memory load stays the same as in Experiment 1. If the effects we found in Experiment 1 were entirely due to differences in working memory load, the results of Experiment 2 should be the same as those of Experiment 1. If, however, the smaller and delayed repetition effect in continuous speech shown by the foreign listeners is mainly due to their segmentation difficulties, the difference in repetition effect between native and non-native listeners should be reduced in the second experiment.

Experiment 2: Test Phase

For the Sentence Test Phase, Figure 3 shows the ERPs to the first familiarized word and the first and second presentation of the unfamiliarized word in the sentences, for the two groups separately. Comparison of Figure 3 with Figure 2 reveals that the repetition effect size in Experiment 2 is somewhat reduced and delayed compared with Experiment 1. Importantly, however, in Experiment 2 the size of the repetition effect for familiarized words in continuous speech did not differ between native and foreign listeners. In contrast to Experiment 1, in this experiment there was *no Repetition by Group interaction* in the 400-900 ms time window for the repetition effect to the familiarized words encountered in continuous speech ($F(1,22)=2.65$, $p=.118$, *Supporting Table 5a, Appendix 2D*). A main effect of Repetition was observed ($F(1,22)=13.57$, $p=.001$). The Repetition effect lasted from 465-910 ms (*Supporting Table 6a, Appendix 2D*). An analysis in this time window (465-910 ms) again failed to show a significant *Repetition by Group* interaction ($F(1,22)=1.67$, $p=.210$).

For novel word repetition within Test (second unfamiliarized – first unfamiliarized) there was no Repetition by Group interaction in the 400-900 ms time window ($F(1,22)=2.51$, $p=.128$, *Supporting Table 5b, Appendix 2D).* However, results of the onset and duration analysis using cluster randomization indicated a Repetition by Group interaction from 600-795 ms (*Supporting Table*

6b, Appendix 2D). The Repetition effect lasted from 600-1090 ms for native listeners, while there was no significant cluster for the foreign listeners.

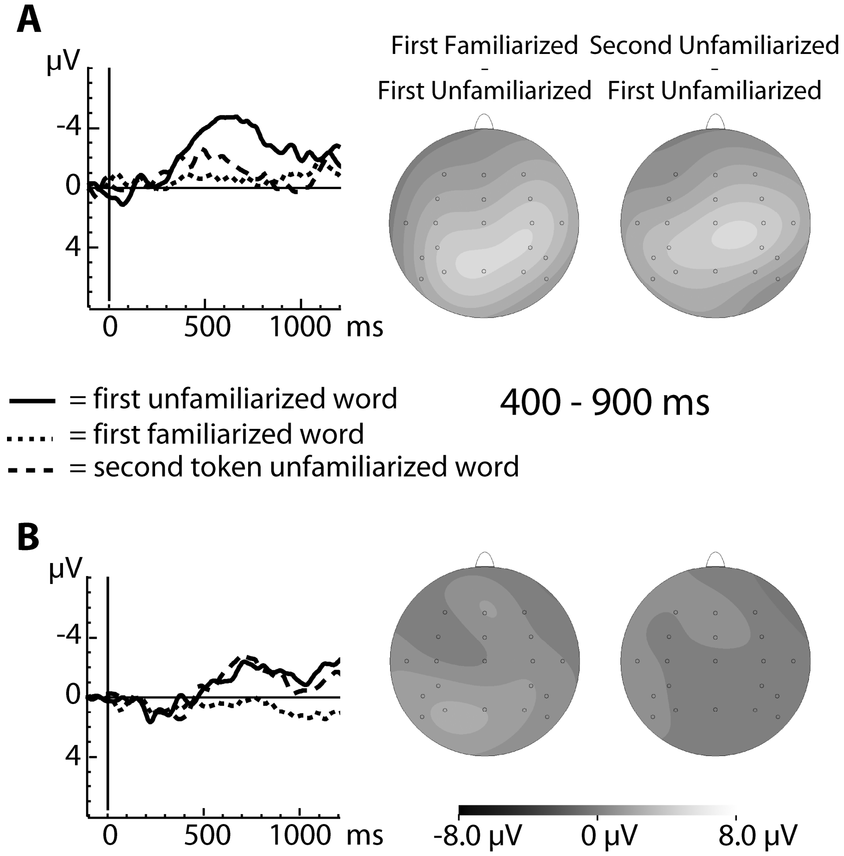

Figure 3: Test phase, Experiment 2*. Repetition Effect in the Sentence Test phase for native (A) and foreign listeners (B). Left: Event-related potential (ERP) from the first familiarized word, and the first and the second occurrence of the unfamiliarized control word in the sentences at a representative electrode site (Pz). Negativity is plotted upwards. Middle and Right: Topographic isovoltage maps of the different repetition effects in the 400 - 900 ms latency range. Middle: recognition of familiarized words in continuous speech: familiarized - unfamiliarized. Right: repetition effects within continuous speech: second unfamiliarized - first unfamiliarized.*

In contrast to Experiment 1, in Experiment 2 the ERP repetition response to familiarized words repeated in sentences did not differ significantly for native and foreign listeners. This suggests that a foreign listener's difficulty in detecting familiarized word forms in the continuous speech signal of an unfamiliar language is indeed at least in part due to segmentation difficulties, and not just to a difference in working memory load induced by foreign rather than native input. However, for novel words repeated in continuous speech the difference in repetition effect between the native and foreign listeners was not abolished. The 100 ms pauses between words were not enough for the foreign listeners to detect the novel word repetition within Test. Thus, the speech segmentation difficulties that foreign listeners encounter cannot be the only reason for the absence of a repetition effect for words repeated within continuous speech. For the familiarized words a memory trace is formed, resulting in successful recognition when word boundaries are made clearer. But the novel unfamiliarized words will have to compete for a place in short term memory with all other words in the sentence (none of them evoking a lexical response). This makes the recognition process extremely difficult for foreign listeners even if the segmentation process is facilitated by inserting pauses between words.

The smaller size and the shorter duration of the repetition effects in Experiment 2 (compared to Experiment 1, see *Supporting Tables 4 and 6, Appendix 2D*) might have multiple origins. First, the absence of coarticulation in Experiment 2 might explain why the repetition effect in this experiment started much later (for natives) than in Experiment 1. Second, the smaller effect sizes in both groups might be the result of an overall signal-to-noise reduction in Experiment 2, due to the absence of a normal intonation contour. Because the materials were constructed by concatenating words recorded in isolation, the sentences lacked a normal intonation contour, and, presumably showed compared to words spoken as part of a sentence, an abnormal phonological variability. As a result, the intelligibility of the speech is likely to have been somewhat reduced. This would make it more difficult for both native and foreign listeners to recognize the repeated words, resulting in later (for natives) and smaller repetition effects. In this way the advantage of the short 100 ms pauses, making segmentation easy (or even redundant), could have been partly counteracted by a loss of intelligibility due to the absence of a normal intonation contour. Nevertheless the differential effect of Familiarization for native and foreign listeners was less pronounced in Experiment 2 than in Experiment 1, and failed to reach significance. This argues against any claim that the effect in Experiment 1 was solely due to a difference in working memory load in native versus foreign listening.

CONCLUSIONS

The ERP repetition effect for words that are repeated in continuous speech is quite different for native and foreign listeners. Even though Dutch and English are highly similar languages, the neurophysiological evidence presented here shows fast segmentation and recognition by Dutch adults, but a reduced and delayed response for English adults. That is, only the native listeners are able to perform fast segmentation of Dutch sentences. Segmentation of continuous speech is a process which listeners have optimized for application to their native language, with the result that this process becomes a demanding one for foreign listeners. Foreign listeners also cannot call on lexical knowledge (in memory) to find boundaries in the speech stream. The resulting speech segmentation difficulty forms an important part of why understanding a spoken foreign language can be so problematic. The frequently reported subjective impression that speakers of other languages talk extremely fast may be grounded in the brain response delays which we have observed here.

REFERENCES

Cutler, A., 2000-2001. Listening to a second language through the ears of a first. Interpreting. 5, 1-23.

Cutler, A., Butterfield, S., 1992. Rhythmic cues to speech segmentation: Evidence from juncture misperception. Journal of Memory and Language. 31, 218-236.

Cutler, A., Carter, D. M., 1987. The predominance of strong initial syllables in the English vocabulary. Computer Speech & Language. 2, 133-142.

Cutler, A., Mehler, J., Norris, D., Segui, J., 1983. A language specific comprehension strategy. Nature. 304, 159-160.

Cutler, A., Mehler, J., Norris, D., Segui, J., 1986. The syllable's differing role in the segmentation of French and English. Journal of Memory and Language. 25, 385-400.

Cutler, A., Mehler, J., Norris, D., Segui, J., 1989. Limits on bilingualism. Nature. 340, 229-230.

Cutler, A., Norris, D. G., 1988. The role of strong syllables in segmentation for lexical access. Journal of Experimental Psychology: Human Perception and Performance. 14, 113-121.

Cutler, A., Otake, T., 1994. Mora or phoneme - further evidence for language-specific listening. Journal of Memory and Language. 33, 824-844.

Dumay, N., Frauenfelder, U. H., Content, A., 2002. The role of the syllable in lexical segmentation in French: Word-spotting data. Brain and Language. 81, 144-161.

Jusczyk, P. W., 1999. How infants begin to extract words from fluent speech. Trends in Cognitive Sciences. 3, 323-328.

Jusczyk, P. W., Aslin, R. N., 1995. Infants' detection of the sound patterns of words in fluent speech. Cognitive Psychology. 29, 1-23.

Kolinksy, R., Morais, J., Cluytens, M., 1995. Intermediate representations in spoken word recognition: Evidence from word illusions. Journal of Memory and Language. 34, 19-40.

Kooijman, V., Hagoort, P., Cutler, A., 2005. Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. Cognitive Brain Research. 24, 109-116.

Maris, E., 2004. Randomization tests for ERP-topographies and whole spatiotemporal data matrices. Psychophysiology. 41, 142-151.

Otake, T., Hatano, G., Cutler, A., Mehler, J., 1993. Mora or syllable? Speech segmentation in Japanese. Journal of Memory and Language. 32, 258-278.

Pfitzinger, H. R., Tamashima, M., Comparing perceptual local speech rate of German and Japanese speech. Proc. of the 3rd Int. Conf. on Speech Prosody, Vol. 1, Dresden, 2006, pp. 105-108.

Rugg, M. D., 1985. The effects of semantic priming and word repetition on event-related potentials. Psychophysiology. 22, 642-647.

Rugg, M. D., Doyle, M. C., Event-related potentials and stimulus repetition in direct and indirect tests of memory. In: H.-J. Heinze, et al., Eds., Cognitive Electrophysiology. vol. Birkhäuser, Boston, 1994, pp. 124-148.

Rugg, M. D., Doyle, M. C., Wells, T., 1995. Word and nonword repetition within-modality and across-modality - an event-related potential study. Journal of Cognitive Neuroscience. 7, 209-227.

Sanders, L. D., Neville, H. J., 2003a. An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. Cognitive Brain Research. 15, 228-240.

Sanders, L. D., Neville, H. J., 2003b. An ERP study of continuous speech processing: II. Segmentation, semantics, and syntax in non-native speakers. Cognitive Brain Research. 15, 214-227.

Sanders, L. D., Newport, E. L., Neville, H. J., 2002. Segmenting nonsense: an event related potential index of perceived onsets in continuous speech. Nature Neuroscience. 5, 700-703.

Suomi, K., McQueen, J. M., Cutler, A., 1997. Vowel harmony and speech segmentation in Finnish. Journal of Memory and Language. 36, 422-444.

Takashima, A., Jensen, O., Oostenveld, R., Maris, E., van de Coevering, M., Fernandez, G., 2006. Successful declarative memory formation is associated with ongoing activity during encoding in a distributed neocortical network related to working memory: A magnetoencephalography study. Neuroscience. 139, 291-297.

Vroomen, J., Tuomainen, J., de Gelder, B., 1998. The roles of word stress and vowel harmony in speech segmentation. Journal of Memory and Language. 38, 133-149.

Vroomen, J., van Zon, M., de Gelder, B., 1996. Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. Memory & Cognition. 24, 744-755.

Summary and conclusions

Chapter 6

Adults listening to their native language are usually unaware of the complexity of listening to speech. Not until they listen to a foreign language do they realize how difficult it is. Recognizing individual words in spoken language becomes an obvious problem in a situation like this. Speakers of a foreign language seem to talk unnervingly fast, and it seems impossible to know where one word ends and the next word begins. In fact, listeners to a foreign language rate that language as faster than native listeners do (Pfitzinger & Tamashima, 2006). Infants, however, are able to recognize some words in their native language even in their first year of life, before they know the meaning of the words. This ability to find word boundaries in spoken language, i.e., word segmentation, has been the topic of this thesis. The main part of the thesis focused on the beginnings of word segmentation in the second half of the first year of life, and the role of metrical stress in the accomplishment of this task (chapters 2, 3, and 4). In addition, word segmentation in native and foreign listeners to Dutch was studied (chapter 5). All studies used an online ERP repetition paradigm to study word segmentation in continuous speech. Converging ERP and behavioral measures were used to study segmentation in the seven-month-old participants (chapter 4).

SUMMARY OF THE RESULTS

The beginning of word segmentation

An important drawback of studies on early word segmentation, until recently, was the lack of an online measure. Only the end result of sentence processing was measurable, for example, with the HPP method. In the past decade, ERP has become a valuable online tool in language research in adults, in particular in studies on sentence processing. In the last couple of years, ERP also has become a more popular tool to study language processing in infants and children. However, word segmentation had not been addressed yet. In chapter 2 of this thesis, the first ERP evidence of word segmentation from continuous speech in Dutch ten-month-olds was presented. Early word segmentation of nouns with a trochaic (strong-weak) stress pattern was studied. In line with the expectations based on the results of behavioral studies using the HPP method (e.g., Jusczyk & Aslin, 1995; Jusczyk, Houston & Newsome, 1999), an ERP effect of word segmentation to the familiarized strong-weak words was found. A clear left-lateralized effect showed that ten-month-olds initiate a segmentation response roughly halfway through the word. Thus, they do not need to hear the entire word to initiate word segmentation.

In chapter 3, the role of strong syllables in weak-strong word segmentation was explored in Dutch ten-month-olds. Although the majority of nouns in Dutch start with a strong syllable, there still are a considerable number of words that start with a weak syllable. Infants at some point have to learn to combine different cues in the language to efficiently segment these iambic (weak-strong) words from speech as well. The metrical stress cue does not suffice to do this. Combining this cue with other cues, such as phonotactic (i.e., possible phoneme order) and phonetic (i.e., properties of speech sounds) regularities in the language, may help to find the word boundaries of iambic words. In general, it is assumed that infants learn to combine different sources of information between eight and ten months of age (e.g., Werker & Tees, 1999). Thus, at ten months of age, some level of weak-strong word segmentation should be possible (e.g., Johnson, 2005). Chapter 3 deals with the ability of Dutch ten-month-olds to segment iambic words from speech, and explored the ERP response to iambic words in isolation and in sentences. The ERP repetition response to isolated words with stress on the second syllable is very similar to that for words with stress on the first syllable. The onset of the repetition response occurs well before the end of the first syllable in both cases. This indicates that, when the word is surrounded by silence, infants process the weak-strong words from word onset. Thus, it is not just the strong syllable infants respond to. In the sentences, however, the ERP response to iambic words is time-locked to the second, strong, syllable and not the first. It appears that Dutch ten-month-olds still largely rely on the strong syllable for word segmentation. In the same experiment, sentences were presented with strong-weak target words with the same strong syllable as in the iambic words. A small ERP response was found time-locked to the strong, first, syllable of the trochaic words. Although the ERP response is smaller and different from the response to the strong syllable in the iambic words, a recognition response was triggered. These results indicate that ten-month-old Dutch infants still strongly rely on the strong syllable for word segmentation. However, the differences between the ERP responses to the iambic and trochaic words, in terms of polarity and distribution, suggest that context does matter. If infants were responding to the strong syllables regardless of context, one would expect to see a similar ERP response for both iambic and trochaic words. This is not what was found.

In chapter 4, converging behavioral and ERP methodologies were used to study word segmentation in Dutch seven-month-olds. Although previous behavioral studies with Dutch infants did not show word segmentation until nine months of age (Kuijpers, Coolen, Houston & Cutler, 1998), the ERP results in this chapter revealed an ERP effect of segmentation already at seven months of age. The HPP study, for which the same materials were used as in the ERP study, did not show evidence of word segmentation. These seemingly conflicting results show that the beginnings of word segmentation do not translate to behavior yet. Even though infants are sensitive to the trochaic stress pattern of their native language, even in continuous speech, this sensitivity is not strong enough to already initiate the corresponding differential head turn. This study reveals the strengths and weaknesses of both the HPP and ERP studies. HPP studies may not be able to pick up on the very early sensitivities to language cues, but are highly valuable as a tool to study the behavioral outcome of changes in the brain. ERP on the other hand can pick up on learning processes not visible as behavior yet (also see McLaughlin, Osterhout & Kim, 2004). However, the relationship between changes in the brain and behavior requires more than just ERP. Converging behavioral and brain measures are necessary to study this relationship.

Listening to native and foreign languages

In adults, lexical knowledge of the native language, in addition to knowledge about pre-lexical cues, combine to efficiently segment words from speech. Foreign listeners obviously lack the knowledge of these language-specific cues. This makes word segmentation particularly difficult for foreign listeners who only command a few words of the language. The experiment reported in chapter 5 explored word segmentation in both native and foreign adults listening to Dutch sentences. Comparable to the ERP studies by Rugg, Doyle and Wells (1995) on word repetition, a positive repetition response was found for words presented in sentences after familiarization in isolation. Both the Dutch listeners and the foreign listeners without any knowledge of Dutch showed this repetition response. However, in spite of highly similar ERP responses to Dutch words presented in isolation, the Dutch listeners showed a very early repetition effect with an onset of 115 ms whereas the foreign listeners showed this response only 515 ms after word onset. This delay in word segmentation may explain the often reported overestimate of the pace with which non-native languages are spoken.

The considerable delay in finding words in continuous speech of a foreign language might make it hard to keep up with the speaker.

CONCLUSIONS

The results of the three experiments on the development of word segmentation show that ERP is a valuable tool to study word segmentation in the first year of life. They provide new perspectives on language development and, in combination with behavioral measures, on the interaction between brain and behavior. However, although the studies were designed to be comparable, especially the studies in Chapters 2 and 4, direct comparisons were not possible due to the different time windows chosen for data analyses. The studies presented in this dissertation were the first to address the seven- and ten-month-old infants' brain response to continuous speech. Therefore, predefined time windows were not available, and visual inspection of the data was required to identify the relevant time windows for each study. Further research is needed to describe infants' brain response to the different cues to word segmentation in more detail and to further define the different ERP components involved. This will improve the comparability of future studies. More generally, it is necessary to explore the development of the ERP signature to language processing early in life in order to get a full understanding of the use of ERP as a tool to study language development.

In the studies presented in this thesis, the focus was on the role of metrical stress in word segmentation. Infants were familiarized with isolated words and tested on familiar and unfamiliar words in sentences. However, in everyday life infants hear words not only in isolation but also in longer utterances. In fact, the majority of the language they hear comes from continuous speech (Van de Weijer, 1998). A HPP study showed that eight-month-old infants familiarized with words in sentences (i.e., without ever hearing the words in isolation) are also able segment these words from speech (Seidl & Johnson, forthcoming). In a follow-up ERP study of the research presented in this thesis, ten-month-old infants will be presented with isolated words after hearing the word previously only in one sentence. In other words, if infants hear a word in a sentence only once, are they able to immediately extract it and recognize it in isolation?

The study in chapter 5 of this thesis was one of the first to use ERP as a tool to study word segmentation in adults (for comparison, see Sanders & Neville, 2003a; 2003b). Behavioral studies established that word segmentation is based on language-specific cues (Cutler, 2000-2001), which is held to be what makes it particularly difficult to segment words from a foreign language. This study was the first to show how this delays the brain response that indicates segmentation of words from continuous speech. It would be interesting to see if this delay reduces with more knowledge of the foreign language. Cutler (2000-2001) suggested that after learning language-specific cues in the native language it is very difficult to learn these cues for a foreign language. Nevertheless, it is easier to recognize words in a familiar foreign language than in an unfamiliar foreign language. Is this due to lexical knowledge only, or do other cues play a role after all? Further research is needed to answer these questions.

Although only a small part of language processing was discussed in this thesis, it is a vitally important part. This thesis sheds new light on the development of early word segmentation and the methods by which it may be studied. Nevertheless, this thesis could not have been written without the patience, persistence and creativity of many other researchers in the field of language development before me. For a tale is but half told, when only one person tells it. I, therefore, end with the wish that the results reported in this thesis might inspire others to investigate in more detail one of the most fundamental questions in language research: How infants are able to bootstrap a lexicon out of the continuously varying speech input in their environment.

REFERENCES

Cutler, A. (2000-2001). Listening to a second language through the ears of a first. *Interpreting, 5*, 1-23.

Johnson, E.K. (2005). English-learning infants' representations of word forms with iambic stress. *Infancy, 7*, 99-109.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1–23.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159–207.

Kuijpers, C., Coolen, R., Houston, D., & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. In *Advances in Infancy Research*. Eds. C. Rovee-Collier, L. Lipsitt and H. Hayne. London: Ablex Publishing.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience, 7*, 703-704.

Pfitzinger, H. R., Tamashima, M. (2006). Comparing perceptual local speech rate of German and Japanese speech. *Proc. of the 3rd Int. Conf. on Speech Prosody, Vol. 1*, Dresden, pp. 105-108.

Rugg, M. D., Doyle, M. C., Wells, T. (1995). Word and nonword repetition within-modality and across-modality - an event-related potential study. *Journal of Cognitive Neuroscience, 7*, 209-227.

Sanders, L.D., & Neville, H.J. (2003a). An ERP study of continuous speech processing I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research, 15*, 228-240.

Sanders, L.D., & Neville, H.J. (2003b). An ERP study of continuous speech processing II. Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research, 15*, 214-227.

Seidl, A., & Johnson, E.K. (forthcoming). Utterance boundaries and infants' extraction of words from fluent speech.

Van de Weijer, J. (1998). Language input for word discovery. PhD Dissertation, Radboud University Nijmegen. *MPI Series in Psycholinguistics 9*.

Werker, J.F., & Tees, R.C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509–535.

Appendices

APPENDIX 1: STIMULUS MATERIALS

APPENDIX 1A: STIMULUS MATERIALS OF CHAPTER 2, 4, AND 5

nr.	target words	sentences
1	hommel	De hommel vliegt van bloem naar bloem.
		Het is een oude hommel met gele strepen.
		Een kleine hommel zit op het gordijn.
		Vaak kan een hommel erg hard zoemen.
2	mammoet	De mammoet is lang geleden uitgestorven.
		Er is een oude mammoet in het museum.
		Die kleine mammoet zwemt in de rivier.
		Daar is een mammoet met veel vriendjes.
3	hofnar	De hofnar maakt weer eens rare grappen.
		De koning hoort de boze hofnar vallen.
		Gelukkig vangt de lange hofnar hem nog op.
		Zonder een hofnar lacht er nooit iemand hier.
4	python	De python ziet er nogal gevaarlijk uit.
		Daar zie ik een boze python liggen.
		Dat is een lange python met scherpe tanden.
		Met een python moet je altijd voorzichtig zijn.
5	gondel	Die gondel wordt elk voorjaar weer gebruikt.
		Dat is een gondel van de stevige slager.
		Mario bouwde een grote gondel voor zijn dochter.
		Die nieuwe gondel moet nog geverfd worden.
6	otter	Die otter is dol op spelletjes doen.
		Piet zag een otter uit een ander land.
		Daar ligt een grote otter op een steen.
		Die nieuwe otter vond snel een vriendje.
7	fakir	De fakir loopt zomaar over de kolen.
		Er is een moedige fakir op de kermis.
		Die oude fakir is bevriend met de dwerg.
		Gisteren bezocht nog een andere fakir onze school.
8	poema	De poema kijkt nieuwsgierig naar de tijger.

		Daar loopt een moedige poema uit het circus.
		De oude poema loopt rusteloos door zijn kooi.
		Een bewaker geeft de andere poema te eten.
9	orka	De orka kan heel goed kunstjes leren.
		Een andere orka is te zien in het aquarium.
		Het is een mooie orka met een grote vin.
		Ik zag een orka op de televisie.
10	emoe	De emoe komt vooral voor in Australië.
		Die andere emoe kan wel heel erg snel lopen.
		Daar staat een mooie emoe naast die grote boom.
		Dat is een emoe van de boerderij.
11	zwaluw	Een zwaluw vliegt vaak laag over het landschap.
		De kleine zwaluw kan heel goed vliegjes vangen.
		Ik zie een andere zwaluw in de wei.
		's Ochtends is die zwaluw altijd erg actief.
12	viking	Een viking gaat op reis naar verre landen.
		Die kleine viking is niet sterk maar slim.
		Dat is die andere viking met veel vijanden.
		Pieter zag die viking uit het noorden.
13	serre	Hier in de groene serre kan je zitten.
		Die serre bij het restaurant is mooi.
		Mijn moeder wil ook een serre van glas.
		Oma had een bijzondere serre vol planten.
14	krekel	Ik zag een groene krekel in het gras.
		Die krekel kan aardig wat lawaai maken.
		In dat verhaal speelt een krekel de hoofdrol.
		Dat is een bijzondere krekel uit Zuid-Amerika.
15	drummer	De drummer speelt soms in de stad.
		Daar is de jonge drummer van de band.
		Een bijzondere drummer is moeilijk te vinden.
		Er is een jonge drummer in het café.
16	hinde	De hinde sprong net op tijd weg.
		Er springt een jonge hinde over de sloot.
		De bijzondere hinde rent door het bos.
		Daar eet een hinde het verse gras.

17	klamboe	Onder zo'n klamboe slaap je echt beter.
		In Afrika is een klamboe echt nodig.
		Daar kan je een oude klamboe kopen.
		Die klamboe van mijn ouders is kapot.
18	toffee	Maar zo'n toffee kleeft wel heel erg.
		Ik eet graag een toffee na school.
		Er ligt nog een oude toffee daar.
		Die toffee smaakt heerlijk bij de thee.
19	logo	Het vorige logo van dat bedrijf is niet mooi.
		Zo'n logo heb ik eerder gezien.
		In de folder staat een logo van die stichting.
		Ze schilderen het echte logo op het raam.
20	kajak	De vorige kajak van Klaas is nog wel bruikbaar.
		Zo'n kajak is alleen voor wedstrijden.
		Ik voel me in een kajak niet echt veilig.
		Hij bouwt een echte kajak van dat hout.
21	ketjap	De rode ketjap is meestal extra scherp.
		Jan doet zijn ketjap altijd over de rijst.
		De ketjap staat in dat blauwe kastje.
		Geef mij die nieuwe ketjap eens aan.
22	tabberd	Die rode tabberd staat de Sint goed.
		Hij draagt zijn tabberd altijd in de winter.
		De tabberd hangt nu aan de kapstok.
		Dat is de nieuwe tabberd uit Spanje.
23	kiwi	De kiwi is een rare vogel zonder vleugels.
		Natuurlijk is een kiwi ook een vrucht.
		Die grote kiwi heeft een lange snavel.
		Gisteren zag ik een kleine kiwi in het reservaat.
24	sheriff	De sheriff is erg belangrijk voor het dorp.
		Buiten loopt een sheriff langs het huis.
		Een grote sheriff ziet er indrukwekkend uit.
		Morgen komt er een kleine sheriff naar de filmset.
25	krokus	Ik vind zo'n witte krokus altijd erg mooi.
		In de pot staat een kleine krokus te bloeien.
		Een krokus is ook leuk om kado te geven.

		De roze krokus zie je vaak.
26	slede	Bas heeft een witte slede in de garage.
		Van de berg gaat zo'n kleine slede extra hard.
		Een slede heb je in sommige landen echt nodig.
		Die roze slede is erg opvallend.
27	pelgrim	De oude pelgrim maakt een reis naar Lourdes.
		De pelgrim is blij met de openbaring.
		Dankzij de jonge pelgrim kon de ezel toch mee.
		Met verbazing keek de dikke pelgrim naar het beeld.
28	mosterd	Die oude mosterd smaakt echt niet meer goed.
		De mosterd wordt verkocht bij elke slager.
		Bij de jonge mosterd past een goed stuk kaas.
		Voor soep is de dikke mosterd ook te gebruiken.
29	pudding	Met een pudding als toetje heb je altijd succes.
		Na een warme pudding drink ik graag koffie.
		De pudding is niet goed gelukt.
		Bij de winkel kan je lekkere pudding kopen.
30	sauna	Naast een sauna hebben ze daar ook een zwembad.
		In een warme sauna kan je goed ontspannen.
		De sauna is behoorlijk ver weg.
		Na het sporten is een lekkere sauna heerlijk.
31	tuba	Uit zo'n tuba komt vaak flink wat lawaai.
		De muzikant poetst zijn tuba elke dag.
		De tuba is een erg groot instrument.
		Met een mooie tuba maak je veel indruk.
32	medley	Met zo'n medley kun je altijd goed meezingen.
		De zanger oefent zijn medley al uren.
		De medley hoorde ik op de radio.
		Een hele mooie medley hoor je slechts zelden.
33	sandwich	Op de sandwich zit kaas en ham.
		In het café kan je een sandwich kopen.
		Na zo'n grote sandwich zit je vol.
		Die sandwich ligt al uren in de vitrine.
34	metro	Met de metro ben je sneller thuis.
		Vanuit de stad moet je een metro nemen.

		In een grote metro kunnen veel mensen.
		Die metro is minstens dertig minuten te laat.
35	sitar	Een sitar is een bijzonder maar simpel ding.
		Tegenwoordig zie je de sitar niet zo vaak.
		Op een kleine sitar oefenen is niet moeilijk.
		De bruine sitar is van een beroemde gitarist.
36	knolzwam	Een knolzwam zie je soms in het bos.
		Toch is ook de knolzwam al vrij zeldzaam.
		In een kleine knolzwam zit soms een kaboutertje.
		Die bruine knolzwam staat leuk in een bloemstukje.
37	maestro	De maestro viel bijna van zijn stoel van verbazing.
		Het is de dikke maestro uit Italië.
		Met de grijze maestro kan je altijd goed praten.
		De andere maestro is een nogal druk mannetje.
38	parka	De parka is vooral lekker warm in het najaar.
		Ik draag een dikke parka van wol.
		Ook die grijze parka geef ik aan mijn nichtje.
		Die andere parka kan ik nog wel aan.
39	monnik	De monnik wiedt zijn tuintje dagelijks.
		De strenge monnik draagt een zware habijt.
		Peter ziet de vriendelijke monnik in het hofje.
		Elke week plukt de jonge monnik verse appels.
40	sultan	De sultan bestuurt het kleine landje.
		De strenge sultan regeert met straffe hand.
		Omar geeft de vriendelijke sultan nog een sigaar.
		Volgend jaar komt de jonge sultan naar Nederland.

APPENDIX 1B: STIMULUS MATERIALS OF CHAPTER 3

WS = target words with a weak-strong stress pattern; SW = target words with a strong-weak stress pattern; *pseudowords* are in italic.

	word pairs	**sentences**	
nr.	WS - SW	WS	SW
1	gebroed - broedsel	Het gebroed loopt daar.	Het broedsel vliegt weg.
		Het jonge gebroed hangt rond.	Het jonge broedsel komt uit.
2	getij - tijger	Het wilde getij bedaard.	De wilde tijger springt.
		Na het vrij rustige getij volgt storm.	Het lijkt een rustige tijger te zijn.
3	geruim - ruimte	Veel geruim kost tijd.	Veel ruimte is er niet.
		De baas neemt geruim de tijd.	Zijn buro neemt ruimte in beslag.
4	gekruid - kruidig	Tante wil graag gekruid voedsel.	Vader lust graag kruidig eten.
		Het wordt een erg gekruid gerecht.	Dat was een erg kruidig drankje.
5	verraad - raadsel	Er is groot verraad gepleegd.	Met een groot raadsel zitten.
		Het verraad is doorzien.	Het raadsel is opgelost.
6	vertrek - trekker	Het kleine vertrek is vol.	De kleine trekker doet het.
		In het grote rode vertrek ligt tapijt.	Op de kleine rode trekker zit iemand.
7	verguld - gulden	Dat lijkt wel zo'n verguld metaal.	Het mes heeft zo'n gulden gloed.
		Een zwaar verguld beeld valt.	Zo'n zwaar gulden zwaard roest.
8	verwoed - woedend	Hij doet verwoed zijn best.	Zij doet woedend haar beklag.
		Heel verwoed zoekt hij zijn boek.	Heel woedend holt hij naar huis.
9	terras - raster	Het moet een mooi terras zijn.	Daar is een mooi raster geplaatst.
		Het terras lijkt ruim.	Het raster ligt thuis.
10	terecht - rechter	Het sleuteltje is terecht gekomen.	Het stoepje is rechter gelegd.
		Hij was terecht boos.	Het was rechter dan eerst.
11	tegoed - goedig	Er staat ook geen tegoed open.	Het is echt geen goedig mens.
		Hij doet zich tegoed aan snoep.	Hij stelt zich goedig voor.

12	tekort - korter	Een tekort wordt aangevuld. Met zo'n tekort ben je niet blij.	Een korter stuk wordt geplaatst. Bij zo'n korter touw kan je ook.
13	sedan - danser	De sedan rijdt toch goed. Hij heeft de mooie sedan verkocht.	De danser doet zijn best. Weer is de mooie danser laat.
14	seleen - lener	Het oude seleen zit in buisjes. Dat beetje nieuwe seleen ligt op.	De oude lener betaalt zijn schuld. Die jonge nieuwe lener ziet het.
15	sekuur - kuren	Hij deed sekuur zijn werk. Die sekuur gemaakte soep is lekker	Ze deed kuren bij haar. Die kuren zijn heel erg gezond.
16	sering - ringen	Ze ziet de gewone sering bloeien. Ook die roze sering geurt.	Hij heeft die gewone ringen gekocht. Die kleine roze ringen glanzen.
17	regie - gieter	De zware regie is moeilijk. De regie valt tegen.	De zware gieter staat buiten. De gieter ligt binnen.
18	rebel - beller	Die rebel schreeuwt hard. Die woeste rebel rent weg.	Die beller rijdt verkeerd. De woeste beller spreekt luid.
19	*ressert* - serre	Zij ziet een *ressert* liggen. Hij vangt het glazen *ressert* op.	Hij heeft een serre gemaakt. Ze tekent een glazen serre erbij.
20	refrein - *freinsel*	Hij heeft een aardig refrein gemaakt. Hij zingt dat refrein snel.	Dat is een aardig *freinsel* geworden. Jan gooit dat *freinsel* weg.
21	beloop - loper	Hij zal het op zijn beloop laten. Hij volgt het grillige beloop grondig.	Hij doet snel met zijn loper open. Ze ziet de grillige loper liggen.
22	beleid - leidster	Het nare beleid geeft onrust. Het nogal strenge beleid heeft effect.	De nare leidster gaat weg. De erg strenge leidster geeft op.
23	belast - lastig	Zij is keer op keer belast daarmee. Zijn vader wordt belast door hem.	Het is weer een keer lastig werk. Die klus wordt lastig voor haar.
24	beschut - schutter	Op een beschut plekje zit je fijn. Dat is een goed beschut pleintje.	Aan een schutter gaf hij melk. Hij is een goed schutter

			geworden.
25	pedaal - daalder	Het zeer antieke pedaal gaat stuk.	De nogal antieke daalder is kwijt.
		Het losse pedaal ligt boven.	Die losse daalder vond ik thuis.
26	perron - ronde	Op het lange perron zit niemand.	De zeer lange ronde was moeilijk.
		Het klassieke perron trekt kijkers.	De klassieke ronde is populair.
27	*penar* - narrig	Een *penar* mens loopt langs.	Een narrig gevoel slaat toe.
		Hij heeft een vrij *penar* idee.	Dat is een vrij narrig bericht.
28	pedant - *dantel*	Het is een erg pedant mannetje.	Dat lijkt een erg *dantel* beest.
		Zo pedant doet hij altijd.	Zo *dantel* is ze nooit.
29	*megeel* - geler	Dat is *megeel* uit Egypte.	Het is geler dan voorheen.
		Hij legt wat *megeel* in de la.	Ze ziet wat geler dan anders.
30	meloen - *loenend*	Ook haar meloen smaakt raar.	En haar *loenend* kalf is lief.
		Vaak eet hij meloen toe.	Dan kijkt hij *loenend* weg.
31	mekaar - karig	Heel gauw geven we mekaar een hand.	Ook daarom geven we karig geld uit.
		We zijn mekaar nu zat.	Ze zijn karig met woorden.
32	mezelf - zelfde	Ik geef dat mezelf kado.	Ik denk dat zelfde vaak.
		Volgens mezelf geven we dat.	Volgens zelfde regels leven.
33	gelei - leisel	De slappe gelei was lekker.	Het slappe leisel bood houvast.
		De vieze groene gelei moet weg.	Het nieuwe groene leisel werkt niet.
34	genie - nieter	Hij ziet de grijze genie weer.	Ze is die grijze nieter kwijt.
		Een genie is charmant.	Een nieter wordt gebruikt.
35	*gevu* - vuren	De snelle *gevu* doet men goed.	Het snelle vuren was over.
		Na de rappe *gevu* is de borrel.	Bij het rappe vuren gaat iets mis.
36	genant - *nantig*	Zo'n genant verhaal ken ik niet.	Zo'n *nantig* kado doet me goed.
		Dat is een zeer genant gebeuren.	Ze bezoekt een zeer *nantig* feest.
37	legaat - gaatje	Hij vindt dat legaat op de kast.	Ze ziet dat gaatje in de muur.
		Het legaat blijkt niet geldig.	Het gaatje is weer gedicht.
38	legaal -galig	De huid is galig geworden.	De pas is legaal verkregen.
		Het lijkt galig weefsel te zijn.	Ze lijkt legaal bezig te zijn.
39	*lemaal* - malen	Ze heeft een lang *lemaal* gebouwd.	Het moet heel lang malen daarna.

		Een grof *lemaal* maakt herrie.	Het grof malen is nodig.
40	*levant - vanter*	Hij ziet een rijk *vanter* liggen.	Ze hoort een rijk *vanter* zuchten.
		Geen *levant* is heel groen.	Geen *vanter* gaat op zoek.

APPENDIX 2: SUPPORTING TABLES

APPENDIX 2A: SUPPORTING TABLES OF CHAPTER 2

Supporting Table 3 (Ch. 2)

ANOVA on mean ERP amplitude in the 200 to 500 ms latency range for the target words in the Familiarization phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)				
Fam.	1, 27	9.85	1010.23	.004*
Fam. x Qua.	3, 81	6.34	289.57	.002*
ANOVA per Quadrant				
Left Frontal	1, 27	19.45	415.16	.000*
Right Frontal	1, 27	10.84	496.47	.003*
Left Posterior	1, 27	3.19	311.11	.085
Right Posterior	1, 27	.044	501.15	.835

Note. Fam. = Familiarity; Qua. = Quadrant

*p<.05

Supporting Table 4 (Ch. 2)

ANOVA on mean ERP amplitude in the 350 to 500 ms latency range for the target words in the Test phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Electrode (20)				
Fam.	1, 27	2.24	416.45	.146
Fam. x El.	19, 513	1.68	56.16	.088
ANOVA: Familiarity (2) x Hemisphere (2) x Electrode (10)				
Fam. x Hem.	1, 27	5.01	78.31	.034*
ANOVA per Hemisphere				
Left Hem.	1, 27	.249	232.53	.622
Right Hem.	1, 27	4.84	262.24	.037*

Note. Fam. = Familiarity; El. = Electrode; Hem. = Hemisphere

* p < .05

APPENDIX 2B: SUPPORTING TABLES OF CHAPTER 3

Supporting Table 3 (Ch. 3)

ANOVA on mean ERP amplitude in the 200 to 500 ms latency range for the target words in the Familiarization phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)				
Fam.	1, 19	15.1	239.07	.001*
Fam. x Qua.	3, 57	34.5	27.01	.000*
ANOVA per Quadrant				
Left Frontal	1, 19	16.1	79.09	.001*
Right Frontal	1, 19	14.7	75.10	.001*
Left Posterior	1, 19	5.6	89.13	.028*
Right Posterior	1, 19	12.2	68.07	.002*

Note. Fam. = Familiarity; Qua. = Quadrant

* $p < .05$

Supporting Table 4 (Ch. 3)

Results of the Test Phase: weak-strong target words

Supporting Table 4a (Ch. 3)

ANOVA on mean ERP amplitude in the 680 to 780 ms latency range time-locked to the onset of the first syllables of the weak-strong target words in the Test phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)				
Fam.	1, 19	3.41	765.78	.080
Fam. x Qua.	3, 57	0.77	94.11	.496

Note. Fam. = Familiarity; Qua. = Quadrant

* p < .05

Supporting Table 4b (Ch. 3)

ANOVA on mean ERP amplitude in the 370 to 500 ms latency range time-locked to the onset of the second syllables of the weak-strong target words in the Test phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)				
Fam.	1, 19	5.00	858.24	.037*
Fam. x Qua.	3, 57	1.64	58.47	.194

Note. Fam. = Familiarity; Qua. = Quadrant

* p < .05

Supporting Table 5 (Ch. 3)

Results of the Test Phase: strong-weak target words

Supporting Table 5a (Ch. 3)

ANOVA on mean ERP amplitude in the 55 to 135 ms latency range time-locked to the first syllables of the strong-weak target words in the Test phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)				
Fam.	1, 19	2.02	219.21	.171
Fam. x Qua.	3, 57	3.07	49.74	.042*
ANOVA per Quadrant				
Left Frontal	1, 19	2.46	114.47	.133
Right Frontal	1, 19	5.56	80.65	.029*
Left Posterior	1, 19	0.28	79.13	.606
Right Posterior	1, 19	0.00	45.45	.939

Note. Fam. = Familiarity; Qua. = Quadrant

* $p < .05$

Supporting Table 5b (Ch. 3)

ANOVA on mean ERP amplitude in the 300 to 500 ms latency range time-locked to the first syllables of the strong-weak target words in the Test phase

source	df	F	MSE	p
ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)				
Fam.	1, 19	0.43	809.10	.520
Fam. x Qua.	3, 57	3.59	62.04	.023*
ANOVA per Quadrant				
Left Frontal	1, 19	1.91	241.68	.184
Right Frontal	1, 19	1.41	246.03	.251
Left Posterior	1, 19	0.22	232.64	.641
Right Posterior	1, 19	0.38	258.22	.543

Note. Fam. = Familiarity; Qua. = Quadrant

* $p < .05$

APPENDIX 2C: SUPPORTING TABLES OF CHAPTER 4

Supporting Table 3 (Ch. 4)

ANOVA on mean ERP amplitude for the target words in the Familiarization phase

	source	df	F	MSE	p
220-320 ms	*ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)*				
	Fam.	1, 27	4.64	529.86	.040*
	Fam. x Qua.	3, 81	1.79	63.22	.167
200-500 ms	*ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)*				
	Fam.	1, 27	3.414	287.05	.076
	Fam. x Qua.	3, 81	2.749	22.99	.050*
	ANOVA per Quadrant				
	Left Frontal	1, 27	6.152	104.96	.020*
	Right Frontal	1, 27	2.872	120.08	.102
	Left Posterior	1, 27	1.399	56.75	.247
	Right Posterior	1, 27	1.305	72.35	.263

Note. Fam. = Familiarity; Qua. = Quadrant

* $p <= .05$

Supporting Table 4 (Ch. 4)

ANOVA on mean ERP amplitude for the target words in the Test phase

	source	df	F	MSE	p
350-450 ms	*ANOVA: Familiarity (2) x Quadrant (4) x Electrode (5)*				
	Fam.	1, 27	0.78	305.94	.388
	Fam. x Qua.	3, 81	4.05	33.85	.018*
	ANOVA per Quadrant				
	Left Frontal	1, 27	0.95	111.75	.337
	Right Frontal	1, 27	3.70	94.23	.065
	Left Posterior	1, 27	0.37	94.51	.551
	Right Posterior	1, 27	0.69	82.38	.413
	ANOVA over subset F4, F8, FC4, FT8				
	Fam.	1, 27	4.28	84.97	.048*
430-530 ms	*ANOVA over subset LTP, CP3, and P3*				
	Fam.	1, 27	4.24	43.93	.049*

Note. Fam. = Familiarity; Qua. = Quadrant

* p < .05

Supporting Table 5 (Ch. 4)

ANOVA on mean looking times in the Test phase of the HPP experiment with Version as between-subjects factor

source	df	F	p
ANOVA: Familiarity (2) x Version (2)			
Fam.	1, 24	1.45	.239
Fam. x Vers.	3, 24	0.79	.551

Note. Fam. = Familiarity; Vers. = Version

APPENDIX 2D: SUPPORTING TABLES OF CHAPTER 5

Supporting Table 2 (Ch. 5)

Familiarization Phase Results (second vs. first isolated word, both experiments)

Table 2a

ANOVAs on mean ERP amplitude in the 400-900 ms after word onset.

Source	*df*	*F*	*MSE*	*p*
Omnibus ANOVA: Natives & Foreign listeners				
Repetition	1,44	74.42	96.55	0.000 *
Rep x Group	1,44	1.22	96.55	0.276
Rep x Experiment	1,44	0.04	96.55	0.852
Rep x Gr x Exp	1,44	0.47	96.55	0.497
Rep x Quadrant	3,132	33.30	16.83	0.000 *
Rep x Qua x Gr	3,132	0.38	16.83	0.677
Rep x Qua x Exp	3,132	1.40	16.83	0.253
Rep x Qua x Gr x Exp	3,132	0.05	16.83	0.952
ANOVA per quadrant				
Left Anterior	1,46	19.92	28.40	0.000 *
Right Anterior	1,46	30.44	26.79	0.000 *
Left Posterior	1,46	91.60	32.94	0.000 *
Right Posterior	1,46	103.01	37.64	0.000 *

* p < .05

Table 2b

Onset analysis results: testing onset and duration of the repetition effects using a cluster randomization procedure (see *Methods*).

Source	*cluster*	*time window (ms)*	*size*	*sumstat*	*p*
Rep. x Experiment	No significant clusters (largest cluster: p=0.391)				
Rep. x Group	No significant clusters (largest cluster: p=0.104)				
Repetition (main effect)	1	240-1200	4494	26644	0.000 *
	2	90-180	313	959	0.039 *

Note: Only significant clusters are given (1=largest cluster, 2=second largest cluster, etc.). "Time window" denotes when the repetition effect is happening (in ms after word onset). "Size" gives the number of (electrode, time point-) pairs included in the cluster, "sumstat" the summed T-statistic of the relevant cluster. The p-value denotes the probability of finding such a large cluster if there was actually no effect.

* $p < .05$

Supporting Table 3 (Ch. 5)

Experiment 1, Sentence Test phase. ANOVAs on mean ERP amplitude in the 400-900 ms after onset of the critical word.

Table 3a

Recognition of familiarized words in continuous speech for natives and foreign listeners: first familiarized vs. first unfamiliarized word (see **Table 1** for example).

Source	*df*	*F*	*MSE*	*p*
ANOVA: Repetition (2) x Group (2) x Quadrant (4) x Electrode (5)				
Repetition	1,22	40.81	38.15	0.000 *
Rep x Group	1,22	4.78	38.15	0.040 *
Rep x Quadrant	3,66	7.68	11.92	0.001 *
Rep x Qua x Gr	3,66	0.86	11.92	0.434
***Natives**: ANOVA:Repetition (2) x Quadrant (4) x Electrode (5)*				
Repetition	1,11	23.95	58.55	0.000 *
Rep x Quadrant	3,33	2.45	18.68	0.113
***Foreign listeners:** ANOVA: Repetition (2) x Quadrant (4) x Electrode (5)*				
Repetition	1,11	18.98	17.76	0.001 *
Rep x Quadrant	3,33	8.55	61.42	0.001 *
***Foreign Listeners**: ANOVA per quadrant*				
Left Anterior	1,11	3.10	6.65	0.106
Right Anterior	1,11	0.54	12.14	0.477
Left Posterior	1,11	28.72	6.86	0.000 *
Right Posterior	1,11	33.25	7.30	0.000 *

* p < .05

Table 3b

Repetition effects within continuous speech for natives and foreign listeners: second vs. first unfamiliarized word.

Source	*df*	*F*	*MSE*	*p*
ANOVA: Repetition (2) x Group (2) x Quadrant (4) x Electrode (5)				
Repetition	1,22	7.39	59.32	0.013 *
Rep x Group	1,22	7.42	59.32	0.012 *
Rep x Quadrant	3,66	1.42	13.62	0.253
Rep x Qua x Gr	3,66	0.11	13.62	0.880
Natives*: ANOVA: Repetition (2) x Quadrant (4) x Electrode (5)*				
Repetition	1,11	11.05	79.53	0.007 *
Rep x Quadrant	3,33	1.25	9.23	0.309
Foreign Listeners*: ANOVA: Repetition (2) x Quadrant (4) x Electrode (5)*				
Repetition	1,11	0.00	39.12	0.996
Rep x Quadrant	3,33	0.42	20.12	0.600

* $p < .05$

Supporting Table 4 (Ch. 5)

Experiment 1, onset analysis results Test Phase. Testing onset and duration of the repetition effects using a cluster randomization procedure (see *Methods*).

Table 4a

Recognition of familiarized words in continuous speech: first familiarized vs. first unfamiliarized word.

Source	*cluster*	*time window (ms)*	*size*	*sumstat*	*P*
Repetition x Group	1	330-595	787	2109	0.002 *
Rep. Natives	1	115-1015	3542	12857	0.000 *
Rep. Foreign listeners	1	695-1160	1351	4359	0.000 *
	2	515-690	313	899	0.028 *

* p < .05

Table 4b

Repetition effects within continuous speech: second vs. first unfamiliarized word.

Source	*Cluster*	*time window (ms)*	*size*	*sumstat*	*pval*
Repetition x Group	1	760-970	588	1687	0.014 *
	2	985-1160	482	1316	0.020 *
	3	460-600	362	944	0.035 *
Rep. Natives	1	420-1085	2231	6907	0.005 *
Rep. Foreign listeners	No significant clusters (largest cluster: p=0.446)				

Note: Only significant clusters are given (1=largest cluster, 2=second largest cluster, etc.). “Time window” denotes when the repetition effect is happening (in ms after word onset). “Size” gives the number of (electrode, time point-) pairs included in the cluster, “sumstat” the summed T-statistic of the relevant cluster. The p-value denotes the probability of finding such a large cluster if there was actually no effect.

* p < .05

Supporting Table 5 (Ch. 5)

Experiment 2, Sentence Test phase. ANOVAs on mean ERP amplitude in the 400-900 ms after onset of the critical word

Table 5a

Recognition of familiarized words in continuous speech for natives and foreign listeners: first familiarized vs. first unfamiliarized word

Source	*df*	*F*	*MSE*	*p*
ANOVA: Repetition (2) x Group (2) x Quadrant (4) x Electrode (5)				
Repetition	1,22	13.57	28.86	0.001 *
Rep x Group	1,22	2.65	28.86	0.118
Rep x Quadrant	3,66	2.30	12.29	0.123
Rep x Qua x Gr	3,66	0.79	12.29	0.439

* $p < .05$

Table 5b

Repetition effects within continuous speech for natives and foreign listeners: second vs. first unfamiliarized word.

Source	*df*	*F*	*MSE*	*p*
ANOVA: Repetition (2) x Group (2) x Quadrant (4) x Electrode (5)				
Repetition	1,22	3.72	58.93	0.067
Rep x Group	1,22	2.51	58.93	0.128
Rep x Quadrant	3,66	0.71	10.11	0.501
Rep x Qua x Gr	3,66	0.99	10.11	0.382

* $p < .05$

Supporting Table 6 (Ch. 6)

Experiment 2, onset analysis results Test phase. Testing onset and duration of the repetition effects using a cluster randomization procedure (see *Methods*).

Table 6a

Recognition of familiarized words in continuous speech for natives and foreign listeners: first familiarized vs. first unfamiliarized word.

Group	*cluster*	*time window (ms)*	*size*	*sumstat*	*p*
Repetition x Group	No significant clusters (largest cluster p=0.173)				
Repetition (main effect)	1	465-910	1166	3468	0.001 *

* p < .05

Table 6b

Repetition effects within continuous speech for natives and foreign listeners: second vs. first unfamiliarized word.

Group	*cluster*	*time window (ms)*	*size*	*sumstat*	*p*
Repetition x Group	1	600-795	397	1139	0.008 *
Rep. Natives	1	600-815	615	2194	0.009 *
	2	840-1090	340	1113	0.022 *
Rep. Foreign Listeners	No significant clusters (largest cluster p=0.636)				

Note: Only significant clusters are given (1=largest cluster, 2=second largest cluster, etc.). “Time window” denotes when the repetition effect is happening (in ms after word onset). “Size” gives the number of (electrode, time point-) pairs included in the cluster, “sumstat” the summed T-statistic of the relevant cluster. The p-value denotes the probability of finding such a large cluster if there was actually no effect.

* p < .05

Samenvatting

Woorden herkennen in gesproken taal lijkt simpel voor volwassenen die naar hun moedertaal luisteren. Als een spreker een zin uitspreekt, hoort de luisteraar zonder moeite de afzonderlijke woorden. Het lijkt of er tussen elke twee woorden een korte pauze is ingelast, vergelijkbaar met de ruimte tussen woorden in geschreven taal. In werkelijkheid is dit echter niet het geval. Luister bijvoorbeeld maar eens naar een vreemde taal. De spreker lijkt heel snel te praten, en het vinden van de afzonderlijke woorden in de klankstroom is bijna niet mogelijk. Dit komt doordat gesproken zinnen, in elke taal, niet uit losse woorden met korte stiltes ertussen bestaan maar uit woorden die aan elkaar geplakt zijn en deels overlappen. In het eind van het ene woord is vaak het begin van het volgende al verwikkeld. Deze overlap tussen woorden wordt *coarticulatie* genoemd. Door deze coarticulatie is het niet eenduidig waar het ene woord eindigt en het volgende woord begint. Echter, volwassenen hebben jarenlang ervaring met het luisteren naar de moedertaal. Ze kennen de klankstructuur van de eigen taal, en ook de betekenis van de woorden. Bovendien weten ze welke woordcombinaties vaak of minder vaak voorkomen, en welke klanken veel of juist weinig met elkaar overlappen als ze na elkaar worden uitgesproken. Deze combinatie van kennis maakt het mogelijk om zonder moeite de afzonderlijke woorden in een zin van elkaar te onderscheiden, ook al is er in werkelijkheid sprake van een stroom van klanken. Dit onderscheiden van woorden in de gesproken taal wordt *woordsegmentatie* genoemd. Bij het luisteren naar een vreemde taal hebben we niet de nodige kennis tot onze beschikking om woorden uit gesproken zinnen te segmenteren. Een vreemde taal heeft een andere klankstructuur en woordbetekenis dan de moedertaal waardoor het heel moeilijk is om de afzonderlijke woorden te *segmenteren* uit de gesproken taal.

Pasgeboren kinderen die voor het eerst hun moedertaal horen, hebben evenmin voldoende kennis van de taal om direct woorden van elkaar te kunnen onderscheiden. Het leren van woordbetekenis en het produceren van de eerste woorden gaat pas aan het eind van het eerste en met name in het tweede levensjaar een grote rol spelen. Toch leren kinderen in het eerste levensjaar al heel veel over hun moedertaal. In de eerste dagen na de geboorte zijn kinderen in staat om de klanken van alle talen van elkaar te onderscheiden. In de loop van de daaropvolgende maanden neemt deze vaardigheid af, maar worden de kinderen steeds beter in het herkennen van de klankstructuren van de eigen taal. Ze leren bijvoorbeeld welke klankcombinaties meer of minder voorkomen, en welke klanken veel aan het begin en het eind van woorden voorkomen, of juist midden in een woord. In het Nederlands komt bijvoorbeeld de klankcombinatie 'sch' voor (zoals in *school*) voor, maar alleen aan het begin van woorden. In andere talen, bijvoorbeeld het Engels, komt deze klankcombinatie helemaal niet voor. Behalve dit soort klankinformatie leren kinderen in de eerste maanden van hun leven ook veel over de klemtoonstructuur van de moedertaal. In het Nederlands bestaat deze voor een groot deel uit woorden die beginnen met een lettergreep met een sterke klemtoon gevolgd door een lettergreep met een zwakke(re) klemtoon, zoals in het woord *tijger*. Woorden met een omgekeerd klemtoonpatroon, zoals het woord *getij* komen daarentegen veel minder voor in het Nederlands.

In de tweede helft van het eerste levensjaar hebben kinderen al zoveel geleerd over de klankstructuur van de moedertaal, dat ze op basis daarvan sommige soorten woorden kunnen herkennen in de gesproken taal, zonder dat ze de betekenis van de woorden weten. Of te wel, de eerste stappen op weg naar woordsegmentatie van de moedertaal worden gelegd. Eén van de belangrijkste aanknopingspunten voor het vinden van woorden in de gesproken zinnen op deze leeftijd is de klemtoonstructuur van de taal, in ieder geval in talen zoals het Nederlands, Engels en Duits. Lettergrepen met een sterke klemtoon vallen meer op in de klankstroom dan lettergrepen met een zwakke klemtoon, en kunnen, in talen met een sterk-zwakke klemtoonstructuur, dus gebruikt worden als aanwijzing voor het begin van een woord. Het leren van woordsegmentatie door kinderen in het eerste levensjaar, en de rol die de klemtoonstructuur van de Nederlandse taal daarbij speelt, is het onderwerp van deze dissertatie.

Het onderzoek naar woordsegmentatie dat besproken wordt in deze dissertatie is uitgevoerd door middel van het meten van de elektrische signalen die door de hersenen geproduceerd worden, of te wel door het meten van een *electroencephalogram* (EEG), terwijl de kindjes naar gesproken taal luisterden. Het meten van een EEG kan informatie geven over de manier waarop de gesproken taal verwerkt wordt in de hersenen. De deelnemer krijgt voor het onderzoek een EEG kapje op het hoofd geplaatst, waarin elektroden zitten. Deze elektroden kunnen aan het hoofd de elektrische signalen oppikken die door de hersenen gegenereerd worden, bijvoorbeeld tijdens het luisteren naar bekende en onbekende woorden. Na afloop van het onderzoek worden de gemiddelden van de hersensignalen berekend voor alle bekende en onbekende woordjes. Deze gemiddelden worden *Event Related Brain Potentials* of *ERP*s (zie figuur 5 van hoofdstuk 1) genoemd. Door het ERP van de bekende woorden te vergelijken met het ERP van de onbekende woorden, kan getest worden of de bekende en onbekende woorden op dezelfde of een andere manier verwerkt worden. Bovendien heeft de ERP techniek een hoge tijdsresolutie (in de orde van grootte van milliseconden), waardoor er heel precies gekeken kan worden hoeveel informatie van een gesproken woord nodig is om een woord te herkennen, zowel als het in isolatie gepresenteerd wordt als wanneer het in een gesproken zin gepresenteerd wordt.

In hoofdstuk 2 en 3 is onderzocht in hoeverre Nederlandse kinderen van tien maanden in staat zijn tot woordsegmentatie op basis van de klankstructuur van de taal (dus zonder dat ze de betekenis van de woorden kennen). Eerdere gedragsstudies wijzen er op dat kinderen rond deze leeftijd dit al kunnen. De studies in deze dissertatie zijn echter de eerste waarbij gebruik gemaakt wordt van het meten van ERP tijdens het luisteren naar zinnen bij kinderen van tien maanden. In het onderzoek dat in hoofdstuk 2 besproken wordt zijn alleen woorden gebruikt die beginnen met een sterke klemtoon, bijvoorbeeld de woorden *hommel, krekel, hinde* en *serre*. Eerst kregen alle kinderen een woordje tien maal te horen. Direct daarna kregen ze acht zinnen te horen, waarvan er vier het eerder gehoorde woordje bevatten, en de andere vier een onbekend woordje met dezelfde klemtoonstructuur (zie Tabel 1 van hoofdstuk 2 voor een voorbeeld. In de Appendices staan alle woordjes en zinnen). Het vergelijken van de ERPs van de bekende en onbekende woordjes liet zien dat de kinderen de bekende woordjes inderdaad al kunnen herkennen in een gesproken zin. Bovendien lieten de kinderen al herkenning zien aan het eind van de eerste lettergreep. Het lijkt er dus op dat kinderen van tien maanden inderdaad de eerste lettergreep met de sterke klemtoon kunnen gebruiken om het begin van een woordje te vinden in de gesproken taal.

In het onderzoek uit hoofdstuk 3 is deze vaardigheid verder onderzocht. Hier hoorden kinderen van tien maanden eerst losse woorden met een minder gebruikelijke klemtoonstructuur in het Nederlands, namelijk woorden met de sterke klemtoon op de tweede lettergreep, zoals *getij*. Alhoewel de overheersende klemtoonstructuur in het Nederlands sterk-zwak is, komen er wel woorden voor met een andere klemtoonstructuur. Dit onderzoek was bedoeld om te kijken in hoeverre kinderen van tien maanden deze woorden al in gesproken taal kunnen herkennen. Na de losse woorden hoorden de kinderen zinnen met het eerder gehoorde woordje, en zinnen met een ander woordje waarbij de sterke lettergreep hetzelfde was, bijvoorbeeld *tijger*. Op deze manier was het mogelijk om te onderzoeken of de kinderen specifiek het eerder gehoorde woordje in de zin terug vinden, of met name de sterke lettergreep *tij*. De ERPs lieten zien dat de kinderen zowel op *tij* in *getij* als op *tij* in *tijger* reageerden. De kinderen lieten geen herkenning zien van de eerste, zwakke, lettergreep van de woordjes, zoals *ge* in *getij*. De lettergreep met de sterke klemtoon speelt dus een hele belangrijke rol bij woordsegmentatie uit zinnen. Echter, de ERPs voor de twee woordsoorten zagen er wel verschillend uit. Zo was het effect voor *tij* in *getij* groter dan voor *tijger,* en had het effect een andere oriëntatie. Deze verschillen suggereren dat er een verschil in verwerking is tussen de twee woordsoorten, ook al is de sterke lettergreep in beide woorden hetzelfde. Wellicht maken de kinderen toch gebruik van de informatie uit de rest van het woord, en horen ze dat *tijger* en *getij* niet hetzelfde woord zijn. Het allereerste begin van de segmentatie van woorden met een afwijkende klemtoonstructuur lijkt hier in gang te zijn gezet. Onderzoek bij iets oudere kinderen zou licht kunnen werpen op de volgende stap in de ontwikkeling, waarbij de kinderen waarschijnlijk niet meer alleen de sterke klemtoon gebruiken om het begin van een woord te vinden, maar ook, door middel van andere aanwijzingen in de taal, het begin van woorden met een zwakke klemtoon kunnen vinden.

Bij het onderzoek in hoofdstuk 4 van deze dissertatie is gekeken naar de vroege woordsegmentatie van woorden die beginnen met een sterke klemtoon, namelijk bij kinderen van zeven maanden. Deze kinderen zijn op dezelfde manier getest als de kinderen van tien maanden in hoofdstuk 2. Naast het ERP onderzoek is er ook een gedragsonderzoek gedaan bij een andere groep kinderen van zeven maanden, waarbij gekeken werd of kinderen van deze leeftijd een voorkeur laten zien voor zinnetjes met eerder gehoorde woorden over zinnetjes met onbekende woorden. De resultaten van het ERP onderzoek lieten zien dat kinderen van zeven maanden ook al in staat zijn tot enige vorm van woordsegmentatie. Ze lieten een herkenningsrespons zien in de ERPs voor de eerder gehoorde woordjes. Deze response zag er wel anders uit dan die van de kinderen van tien maanden. Bovendien lieten de kinderen van zeven maanden in het gedragsonderzoek nog geen voorkeur zien voor de zinnen met eerder gehoorde woorden. Deze combinatie van resultaten geeft aan dat kinderen van zeven maanden helemaal aan het begin staan van het leren van woordsegmentatie. Er gebeurt al wel wat in de hersenen als ze eerder gehoorde woorden in een zin terug horen, maar deze reactie is nog niet sterk genoeg om het bijbehorende gedrag aan te sturen in het gedragsonderzoek. De resultaten van deze studie geven aan dat het meten van ERP een waardevolle bijdrage kan leveren aan het onderzoek naar de vroege taalontwikkeling. Leerprocessen die nog niet goed met gedragsstudies getest kunnen worden bij jonge kinderen, kunnen op deze manier toch in kaart gebracht worden. Door ERP studies met gedragsstudies te combineren kan gekeken worden naar de leercurve van bepaalde stappen in de taalontwikkeling.

In de laatste studie van deze dissertatie, beschreven in hoofdstuk 5, is er gekeken naar woordsegmentatie van het Nederlands door Nederlandstalige volwassenen en Engelstalige volwassenen zonder kennis van de Nederlandse taal. Zoals hierboven beschreven, doet het segmentatieprobleem zich niet alleen voor bij het leren van de moedertaal, maar ook bij het luisteren naar een vreemde taal. Hoeveel vertraging er optreedt bij het segmenteren van woorden tijdens het luisteren naar een zin in een vreemde taal is echter niet eerder onderzocht. Om hier meer duidelijkheid over te krijgen is bij deze studie bij volwassenen dezelfde ERP procedure gebruikt als bij de kinderen van tien maanden uit hoofdstuk 2. De resultaten van deze studie laten zien dan Nederlandstalige volwassenen na 115 ms. na het begin van een woord al ontdekt hebben dat er een nieuw woord begonnen is. Dit is buitengewoon snel, in aanmerking genomen dat de woorden gemiddeld circa 700 ms lang waren. De Engelstalige volwassenen hebben veel meer tijd nodig hiervoor, namelijk circa 500 ms. Dit verschil in tijd wordt waarschijnlijk veroorzaakt door het gebrek aan kennis over de Nederlandse taal. Hierdoor is het voor de Engelstalige deelnemers heel moeilijk om het begin en einde van woorden in een zin te ontdekken. De Nederlanders kunnen daarentegen gebruik maken van coarticulatie in de taal en hebben al heel snel door wanneer het eerder gehoorde woord herhaald wordt in de zin. Bij zinnen waarin een woord voorkomt dat de deelnemers niet eerder hebben gehoord (tijdens het onderzoek) hebben de Nederlanders iets meer tijd nodig om het woord te herkennen, maar lukt het de Engelstaligen helemaal niet meer om het betreffende woord uit de gesproken zin te segmenteren. Het zou interessant zijn om te kijken hoe snel dit verandert tijdens het leren van een vreemde taal.

Concluderend kan gesteld worden dat het meten van ERPs een nuttige techniek is voor onderzoek naar het leren van woordsegmentatie en woordherkenning, ook bij kinderen op jonge leeftijd. Dit soort onderzoek kan nieuwe inzichten geven over de vroege taalontwikkeling bij kinderen in het eerste levensjaar, en, in combinatie met gedragsonderzoek, licht werpen op de relatie tussen hersenen en gedrag.

Curriculum Vitae

Valesca Kooijman werd geboren op 26 januari 1974 te Ede. Na het behalen van haar VWO-diploma aan het Erasmus College te Zoetermeer begon zij de studie Psychologie aan de Vrije Universiteit te Amsterdam. In het kader van haar afstudeerrichting Neuropsychologie liep zij in 1996 een half jaar stage aan het Neuropsychiatric Institute van de Univerisity of California, Los Angeles (UCLA). Hier deed zij onderzoek naar het effect van Prozac op het neuropsychologisch functioneren en EEG van zwaar depressieve patiënten. Na haar afstuderen werkte ze van 1998 tot 2001 als onderzoeksassistent in de Neurocognition of Language group op het Max Planck Instituut te Nijmegen. Hierna begon ze aan haar promotieonderzoek in de Comprehension group op het MPI. Het onderzoek maakte deel uit van het Spinoza-project 'Native and Nonnative Listening' van prof. dr. Anne Cutler. Daarnaast werkte ze gedurende vijf jaar als labmanager van drie EEG- en gedragslabs op het F.C. Donders Centre for Cognitive Neuroimaging onder leiding van prof. dr. Peter Hagoort. Sinds maart 2007 werkt ze als postdoc neuroimaging bij het Top Insitute Food & Nutrition te Wageningen.