Coda perception in adults: ERP evidence

Maria A. Muilwijk

s0856924

Thesis

BA Linguistics

Leiden University

August 2012

Supervisors:

dr. Caroline M.M. Junge

dr. Claartje C. Levelt

Table of contents

INTRODUCTION

Coda omission in productions of infants: the result of incomplete lexical representations?

A word like "ten" can be pronounced in different ways, depending on the phonetic context in which it occurs. It can be pronounced as [tem] as in "te[m] bags" or as [teŋ] in "te[ŋ] gates" (Fitzpatrick & Weeldon, 2002). However, these different pronunciations are not considered to be all different morphemes. On the contrary, at some abstract level they are considered to be the same. In the theory of generative phonology (Chomsky & Halle, 1968) two distinct levels of representations are distinguished: the surface representation and the underlying representation. The underlying level concerns a unique representation of every morpheme. This unique representation is the abstract form of the word or morpheme stored in the mental lexicon before any phonological rules have been applied to it. The surface representation on the other hand is the result of phonological rules applied to the underlying form. Therefore the surface form can been seen as a more or less accurate transcription of what is pronounced by the speaker. Both the underlying representations and surface representations are composed of bundles of distinctive features.

In research to first language acquisition it has often been noticed that when infants start producing words, they often omit coda consonants in their attempt to pronounce meaningful words (see e.g. Smith. 1973). Examples of coda omissions from English are given in (1) below (Menn, 1976).

(1) Coda-omissions in child language (Jacob 1;4/1;5)

      a. hat   /hæt/   [hæ]

      b. nose   /noz/   [do]

In the literature, this type of omission is accounted for in terms of constraints that are active on output forms of young children; that is, only syllable templates consisting of

C(onsonant)V(owel) sequences are allowed in production (see e.g. Menn, 1978). In an Optimality Theoretic framework (OT; Prince & Smolenksy 1993) this can been seen as a NOCODA constraint that is initially high-ranked in the grammar of the child. It is assumed to be this restriction that does not allow CVC sequences in the child's output (Levelt, Schiller & Levelt, 2000). Such theories that follow the OT framework believe that there is a constraint on coda in child-grammar at the level of production. However, there are other possible causes for coda omission in the early developing child's speech. Levelt (2012), for example, hypothesized that infants omit codas, because they have incomplete lexical representations in their developing mental lexicon. Therefore, a lexical representation lacking information on the target coda consonant could be a likely source for coda omissions found in productions of 14-month-old infants. To put it differently, in contrast to the OT, Levelt argues that coda omission which we observe in infant's speech is already absent in the representation in the developing mental lexicon. The mapping between surface structure and underlying form is thought to be affected by a constraint that forbids codas. In order to test this hypothesis, two different methods are suitable: the habituation/dishabituation procedure developed by Werker, Cohen, Lloyd, Casasola, & Stager (1998) and elicitation of the Mismatch Negativity (MMN) (Näätänen, Gaillard & Mäntysalo, 1978), while event related brain potentials are measured.

In the sections below two different experiments are described that use these methods in order shed more light on coda representations in the mental lexicon of 14-month-old infants.

Habituation/dishabituation procedure

Levelt (2012) tested 14- and 18-month-old infants on their perceptual sensitivity to coda omissions in a word learning experiment, using the habituation/dishabituation procedure. In this procedure infants are presented with a novel object, while simultaneously hearing a novel syllable over a loudspeaker. After habituation, the same object is paired with a different

syllable. Looking times of the infants are measured and are compared to the preceding trials. If infants notice the acoustic difference between the stimuli, looking times are expected to be longer in the switch trials than in the same trials. In the habituation phase Levelt (2012) presented 14-month-olds a colorful fantasy animal, which was labeled with an auditory presented novel word: /pat/, /pak/, /pas/ or /pa/. After habituation the infants were tested on their ability to detect a phonetic change in the word paired with the object. Sensitivity to coda omission as well as coda addition was tested by presenting a mispronounced version of the word in the test phase. This was [pa] for /pat/, /pas/, /pak/ and [pat] for /pa/. Her predictions were as follows: "if 14-month-olds indeed build incomplete representations during word learning, parsing CV structures from the acoustic input, they should be tolerant of coda omissions in mispronounced items." (Levelt, 2012, p. 175). The results of this study indicate that 14-month-olds indeed seem to be sensitive to mispronunciations, but only when it involves coda addition. As predicted, coda omission was not noticed. This led Levelt to conclude that 14-month-olds do not phonologically parse information from marked syllabic positions, which leads to incomplete lexical representations. The same procedure with 18-month-old infants on the other hand, revealed an increased sensitivity to coda omissions. This indicates that the sensitivity to coda omissions is a developmental effect.

The Mismatch Negativity

Another way to test the amount of phonetic detail in the mental lexicon is by the mismatch negativity (MMN). The MMN reflects an automatic response of the brain to an infrequent acoustic change in an otherwise regular auditory sequence (Näätänen, Gaillard & Mäntysalo, 1978). It appears as a negative deflection and peaks at 100-250 ms from stimulus onset. The maximal amplitude of the component can be found over frontal and central areas of the scalp (for a recent review see, Garrido, Kilner, Stephan & Friston, 2009).

In eliciting the MMN, the oddball paradigm is widely used. In this paradigm a subject is presented a homogeneous sequence of auditory stimuli, the standards. After a certain amount of standard stimuli, the subject is presented an infrequent sound, which is called the deviant. At the same time the subject's brain can be recorded with electrophysiological techniques such as electro-encephalography (EEG) and magneto-encephalography (MEG). The MMN can be measured for both speech (e.g. consonant-vowel pairs such as /ba/ vs. /da/) and nonspeech stimuli (tones of different frequencies: e.g. 1000 Hz vs. 1200 Hz). The result of comparing the average of the standard stimuli and the deviant stimuli is the MMN. Finding an MMN entails that the brain has picked up the difference between the standard and the deviant stimuli; this means that the standard stimuli is well encoded and that the central auditory system has been able to form a representation of the repetitive auditory stimuli before the deviant stimulus was presented. In an auditory oddball paradigm, the MMN can be elicited without paying attention to the stimuli. It is even found in comatose patients (e.g. Fischer et al., 1999; for a review see Näätänen et al., 2007).

One theory that explains the MMN is the trace-mismatch explanation (Näätänen, 1992): the repetitive sound, i.e. the standard, forms a memory trace in the auditory system. An MMN is then elicited if the deviant stimulus following the standard stimulus does not match the memory trace of the preceding stimuli while it is still active (but see Jääskeläinen et al., 2004; Garrido et al., 2008 for different explanations; for a review, see Garrido, Kilner, Stephan & Friston, 2009; or see Näätänen & Winkler,1999). This memory trace can last for about 5-10 s in young and healthy subjects (e.g. Böttcher-Gandor & Ullperger, 1992).

The MMN paradigm can be used to test the amount of phonetic detail in the infant's mental lexicon, because this paradigm allows for comparing the underlying representation of a sound with the surface form of another incoming sound. The logic is as follows. In speaking a language, words are frequently used. Therefore, representations of the mental lexicon are

stored in the long term memory. In an oddball paradigm the standard stimuli is repeated long

enough to create a central sound representation, which corresponds in part to the long-term

memory traces (Nätäänen, 2001; Cowan, 1999; Eulitz & Lahiri, 2004). Since the entries in

the mental lexicon are by definition long term memory representations, the central sound

representation created by the standard stimuli may convey information about the

phonological representation of the sound in the mental lexicon, which is called in linguistic

terms the underlying representation. The deviant stimulus on the other hand is infrequent and

creates a percept that corresponds in part to phonological features that are extracted from the

acoustic signal. In other words, the percept created by the deviant stimulus corresponds to the

so-called surface form. In this way the MMN is an ideal method to study the difference

between the surface form, which is extracted from the deviant stimulus, and the underlying

representation of a sound, created by the standard (Eulitz & Lahiri, 2004).

Recall the results of Levelt (2012) described above. Based on these findings, Nijveld

designed a second experiment (in prep.) in order to investigate the infant's brain response on

coda omissions. As in the above described study, 14-month-old infants were tested on their

perceptual sensitivity to coda omissions. However, instead of using a word learning

paradigm, infants here listened passively to trains of repeated syllables with occasional

deviant syllables (oddball paradigm), while watching silent movies non-synchronized to the

auditory stimuli. In contrast to Levelt (2012), infants did not learn the meaning of novel

words. Wordform learning was here based at the acoustic level. While infants were listening

to these 'words', their event related brain potentials were measured. In this experiment

consisting of two blocks, infants listened to one condition in which the standard stimuli was

/ba:t/ and the deviant /ba:/ and another condition with the reversed setting; that is /ba:/ as the

standard and /ba:t/ as the deviant. They hypothesized that they would find the same

asymmetry in perception as was found by Levelt (2012) for word learning: a tolerance to

omission of coda consonants, but not to additions. Instead of that, the results showed a response to both stimuli types: it seems to be the case that coda addition as well as coda omission are noticed. Data obtained from this study are therefore not in line with the results of Levelt (2012). Further it must be underlined that an MMN is only observed when codas are added. When codas are omitted the response manifested itself in a slow positive drift. It is rather unclear how these data should be interpreted.

To interpret data in infant research, it is often considered worthwhile to understand how the same experiment works with adults. How do adults respond to coda addition and omission? The aim of the current study is therefore to extend the reported studies by this third experiment: an EEG study with adult subjects using the oddball paradigm, with /ba:/ and /ba:t/ as auditory tokens. Following Levelt (2012), we expect the coda constraints active during the perception of 14-month-old infants to annulled in a later stadium of development, since in Dutch phonology the NOCODA constraint is relatively low-ranked; consequently, in adult's perception of codas, NOCODA constraints are not expected to influence the incoming speech any longer. In short, for adults we expect to find an MMN in both directions, i.e. both when codas are added and when they are omitted.

A second goal of this study is to examine the difference in results of a study using the oddball paradigm with multiple tokens versus single token.[1] We hypothesize that the use of a single token will evoke a clearer MMN, since there is no variation in the standard stimuli and therefore an MMN can only occur when the deviant is presented. On the other hand, the use of multiple tokens introduces acoustic variability and creates more natural speech perception conditions. As a result, the processing system is forced to map the incoming acoustic signals onto more abstract representations (Eulitz & Lahiri, 2004). The disadvantages of using

---

[1] The stimuli and results of these pilot studies are discussed in the appendix. We further discuss our findings in the discussion.

multiple tokens is that we do not exactly know how precise the memory trace, created by the standard stimuli, must be in order to evoke an MMN.

The present study differs from previous research in that it tries to evoke an MMN in coda position. Generally, the MMN is evoked immediately after stimulus onset, e.g. by a change from /da/ to /ga/ or by tones that have different frequencies (for a review on different types of stimuli that elicit an MMN, see Näätänen, 1992). Therefore, until now little attention had been paid to time windows in the offset. Because in this study the standard and deviant sound share the onset we do not know when in time participants will notice the difference between the two types of stimuli. When do they hear that a coda has been added to the stimuli, and how does the brain react to silence when the coda has been deleted? We therefore used different markers in time to denote stimulus onset and the offset of the nucleus.

METHODS

Participants

Eighteen healthy adults aged 17-31 (mean age: 21,8; 12 female) participated in the present EEG experiment. Four of them participated originally in the pilot with multiple token condition, but were added in the analysis of the current experiment, too. Three additional participants were excluded from analysis due to technical problems and five because of insufficient trials that could be averaged in one of the conditions (less than 70 trials). All subjects were native speakers of Dutch, righthanded and had no history of hearing or neurological disorders. Participants were informed orally and in writing about the EEG procedure and gave informed consent prior to the experiment. For their participation they were paid 12 Euros or received four credit points.

Stimuli

The stimuli of this experiment comprised four tokens of the pseudowords /ba:/ (mean duration: 637 ms) and /ba:t/ (mean duration: 860 ms). A female native speaker of Dutch

recorded multiple repetitions of the two stimuli with a Sennheiser mkh 416t microphone.

Digital audio recordings were made in a sound-proof booth on an Intel-based PC, using

Adobe Audition, digitized to disk at a 44.1 kHz sampling rate. The stimuli were edited using

Praat software (Boersma & Weenink 2012). Acoustic analyses on these stimuli are presented

in Table 1 and 2. The matching oscillograms are shown in Figure 1 and 2.

	Duration /b/	Duration vowel	Onset ct (/ba:/)	Duration silence	Onset /t/	Duration /t/	Total duration
Baat 1	176	392	568	86	654	195	849
Baat 2	116	421	537	47	584	215	799
Baat 3	162	484	646	60	706	223	929
Baat 4	110	448	558	60	618	245	863
Average	141	436	577	63	641	220	860

Table 1: acoustic details of /ba:t/ in ms.
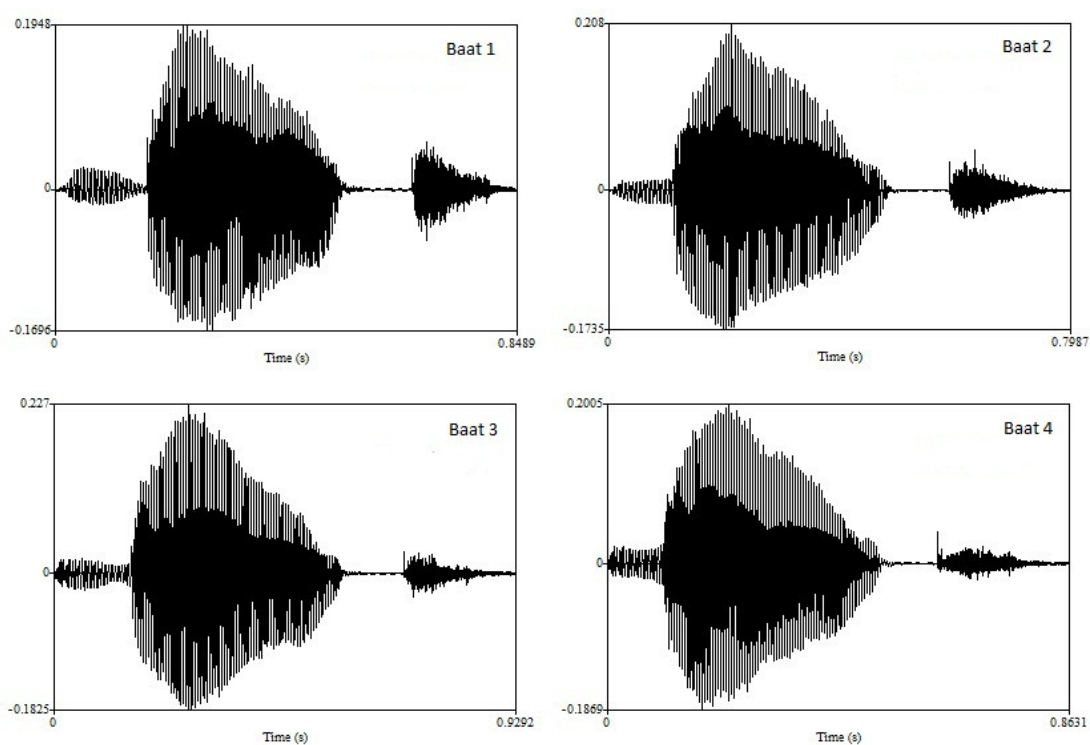
Figure 1: oscillograms of the four /ba:t/ tokens

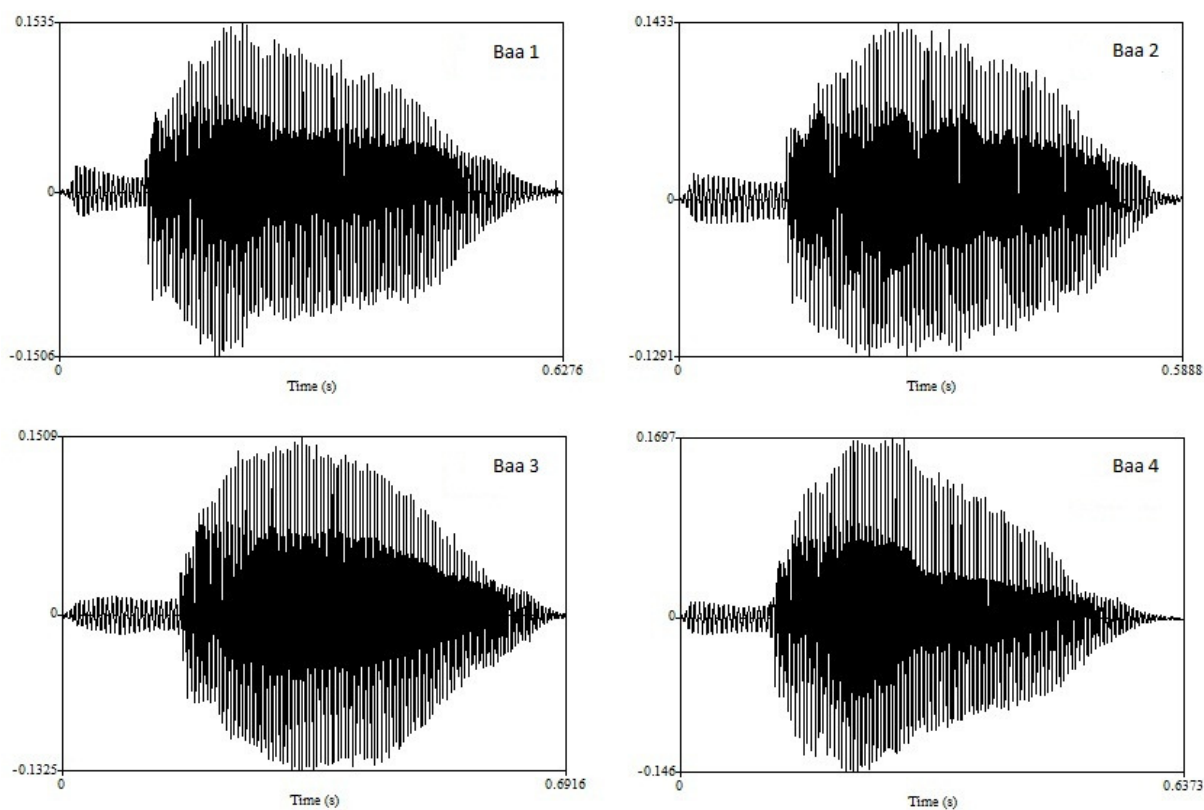|	Duration /b/	Duration vowel	Total duration
Baa 1	117	510	627
Baa 2	128	461	589
Baa 3	168	524	692
Baa 4	126	511	637
Average	135	502	636

Table 2: acoustic details of /ba:/ in ms.

Figure 2: oscillograms of the four /ba:/ tokens

Experimental design

The experiment comprised four blocks, each containing 586 trials consisting of standard (87,2%) and deviant stimuli (12,8%). The mean duration of one block was approximately thirteen minutes. At the beginning of each block, minimal ten standard stimuli were presented

that allowed for a stable regularity representation. Stimuli were pseudo-randomized so that there were at least three standards between any two deviants. When the experiment was paused, we always started with at least four standards, before the experiment continued. Each subject had his own list and all stimuli were presented with an ISI of 500 ms. Two conditions were applied, i.e. /ba:t/ was used as the standard stimulus and /ba:/ as the deviant in condition 1 and the reversed combination, with /ba:/ as standard and /ba:t/ as deviant, was implemented in condition 2, so that both pseudowords served as standard and as deviant. Order of presentation of the different conditions was counterbalanced across participants.

Experimental procedure

The present experiment followed mainly the same experimental design as the study of Nijveld (in prep), but differentiated from that study in the number of blocks (four blocks in the present study instead of two). Participants were instructed to watch the BBC documentary of Madagascar "Land of Heat and Dust" without sound while acoustic stimuli were presented binaurally over loudspeakers placed on the table in front of them. Between blocks, subjects were given a break of self-determined duration.

EEG recordings and data analysis

EEG was recorded using a BioSemi ActiveTwo system (BioSemi, Amsterdam, The Netherlands) with a sampling rate of 512 Hz and a bandpass of 0,06-100 Hz. Recordings were taken from 32 scalp electrodes attached to an electrode cap and were placed according to the 10/20 system (Fp1, AF3, F7, F3, FC1, FC5, T7, C3, CP1, CP5, P7, P3, Pz, PO3, O1, Oz, O2, PO4, P4, P8, CP6, CP2, C4, T8, FC6, FC2, F4, F8, AF4, Fp2, Fz, Cz). Two additional electrodes were placed at the left and right mastoids. To control for blinks and horizontal and vertical eye movements an electrooculogram (EOG) was recorded using four electrodes: two were placed below and above the participant's left eye, one was placed to the left of the left eye, and one to the right of the right eye. The ground electrode during

acquisition was formed by the Common Mode Sense active electrode and the Driven Right Leg passive electrode. Markers were sent on-line to the EEG computer denoting the onset of the stimuli or the moment in time in which participants could hear that the stimuli consisted of a coda /t/. This means that for each of the four / ba:t/ stimuli we had a marker signaling its onset from the /b/ and a marker signaling the offset of the nucleus. Because we used different tokens, the timing between the two markers varied accordingly (range: 537-646 ms, cf. Tabel 1). For the /ba:/ stimuli, we also had two types of markers; one corresponding to stimulus onset, and one corresponding to the offset of the stimulus (note that this is more or less the same time window as the offset of the nucleus for the /ba:t/ stimuli; range: 589-692, cf. Tabel 2).

Off-line analysis was performed using Brain Vision Analyzer software (Brain Products, Gilching, Germany). The EEG was segmented for each trial, beginning 200 ms before each marker and continuing for 800 ms (total: 1000 ms). Epochs with voltage variation exceeding 150 μV at any EEG channel were rejected. Subjects were excluded from analysis if less than 70 trials in one of the conditions could be averaged. To avoid effects due to differences in signal-to-noise ratio, only the standard stimuli presented just before a deviant were included in the analysis, such that the number of accepted EEG epochs did not differ between deviant and standard stimuli for each experimental stimulus. Offline we referenced the EEG signal to linked mastoids and we applied a .1-30 Hz filter. A baseline correction was performed for each individual and each condition by subtracting the average scalp distribution during a 200 ms epoch prior to marker onset.

Statistical analysis

In analyzing the data we compared ERPs from standard stimuli to those of deviant stimuli both for /ba:t/ stimuli as well as for /ba:/ stimuli. Because we did not know when the adult brain would pick up on the difference between standard and deviant stimuli, we used visual

analysis of the waveforms to select time windows of interest as well as the standard 100-250 ms time window suggested by Näätänen (2001). Time windows were selected both in stimulus onset as in stimulus offset. We carried out repeated measures of ANOVA (Analysis of Variance), with the following within-subject factors: Condition (2: standard, deviant), Region (2: anterior, posterior), Hemisphere (2: left, right) and Electrode (7). This created four quadrants of the brain: left anterior (FP1, AF3, F7, F3, FC1, FC5, C3), left posterior (T7, CP1, CP5, P7, P3, PO3, O1), right anterior (FP2, AF4, F8, F4, FC2, FC6, C4), right posterior (T8, CP2, CP6, P8, P4, PO4, O2). For sake of clarity, we only report main effects or interactions with the factor Condition. When we observed significant interactions, we carried out separate ANOVAs for subset of electrodes. For all analyses we used the Huynh-Feldt epsilon correction. Below, the original degrees of freedom are reported along with adjusted p-values and adjusted effect sizes (partial eta-squared: ŋ2).

RESULTS

The grand averages of standard and deviant stimuli at Cz are displayed in Figure 3 (/ba:t/ stimuli) and Figure 4 (/ba:/ stimuli), and can be found in the sections below. In what follows, we describe statistical analysis for every time window that has been selected. We analyzed both the onset of the stimuli and the offset.

ERPs time-locked to the onset of /ba:t/ stimuli

As can be seen from Figure 3a, the standard and deviant waveform start to diverge from each other around 200 ms after stimulus onset. Based on visual inspection we then see a slightly larger difference between the standard and deviant waveform in the 250-350 interval. By the end of the vowel we observed again a deviation of the deviant waveform; that is, in the 450-550 time-window (cf. Table 1 for the average of the duration of the vowel in /ba:t/ stimuli). We therefore chose these time windows for further statistical analysis.
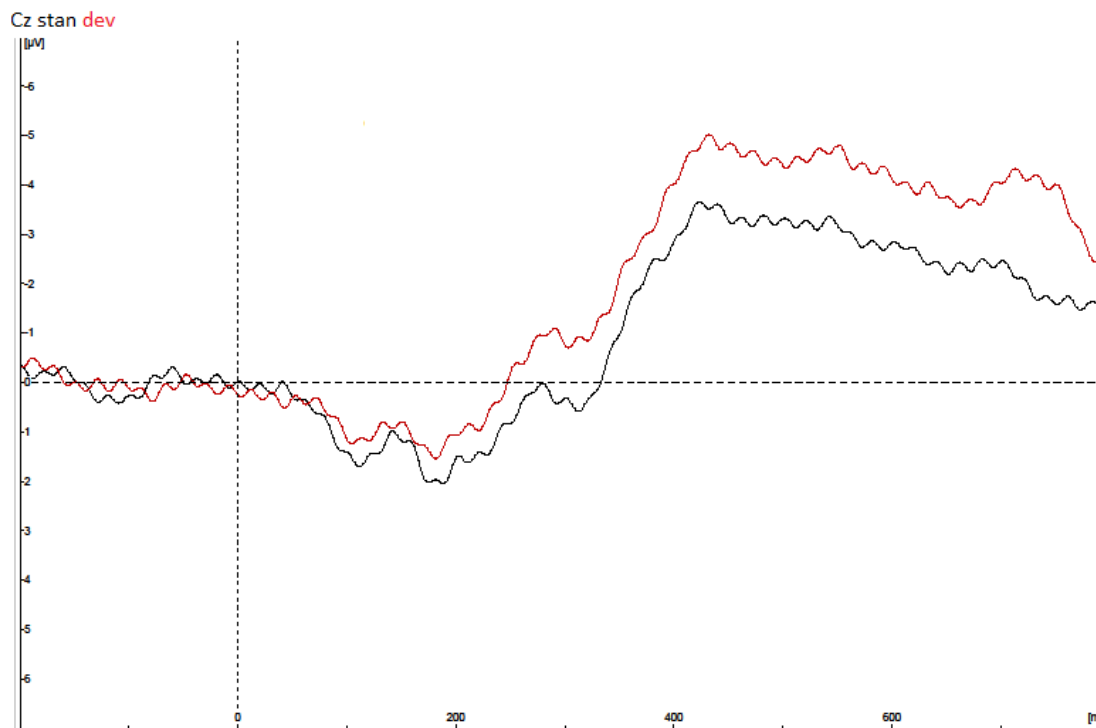
Figure 3a: Grand average (n=18) ERP waveforms at Cz time-locked to the onset of /baːt/.

250-350 ms interval. The peak that we observed during visual inspection in the 250-350 ms interval turned out indeed to be significant. We found a main effect of Condition ($F_{1.17} = 8,172$, $p = 0,011$, $\eta^2 = 0,325$) with a two-way interaction between Condition and Hemisphere ($F_{1.17} = 5,008$, $p = 0,039$, $\eta^2 = 0,228$). We therefore tested for topographic distribution by taking together the electrodes of the left anterior quadrant with the left posterior quadrant and electrodes of the right anterior quadrant with the right posterior quadrant. This created two halves of the brain: the left hemisphere and the right hemisphere. Results of such an analysis showed a main effect of condition in both hemispheres. However, it must be emphasized that the level of significance is much higher in the right hemisphere ($F_{1.17} = 11,974$, $p = 0,003$, $\eta^2 = 0,413$) than in the left hemisphere, which is rather on the border of significance ($F_{1.17} = 4,504$, $p = 0,049$, $\eta^2 = 0,209$).

450-550 ms interval. Statistical analyses showed a main effect of Condition in the 450-550 ms interval, which indicates the end of the vowel ($F_{1.17} = 7,998$, $p = 0,012$, $\eta^2 = 0,320$). Again we found an interaction between Condition and Hemisphere. As we did for the

250-350 ms interval, we tested topographic distribution by performing further post-hoc analysis of the interaction between Condition and Hemisphere, dividing the brain in a left hemisphere and a right hemisphere. Analysis revealed a main effect of Condition in both hemispheres. Again, the level of significance is clearly higher in the right hemisphere ($F_{1.17}$ = 10,591, p = 0,005, η^2 = 0,384) than for the left hemisphere ($F_{1.17}$ = 4,553, p = 0,048, η^2 = 0,211).

ERPs to /ba:t/ stimuli time-locked to the offset of the nucleus

In analyzing ERPs to /ba:t/ stimuli time-locked to the offset of the nucleus, we chose the 100-250 ms interval, which is a standard interval for the MMN negativity (Näätänen, 2001). We further choose two additional time windows, bases on visual analysis. As can be observed in Figure 3b below, two early negative peaks show up in the 50-100 ms and 120-220 ms interval, respectively.
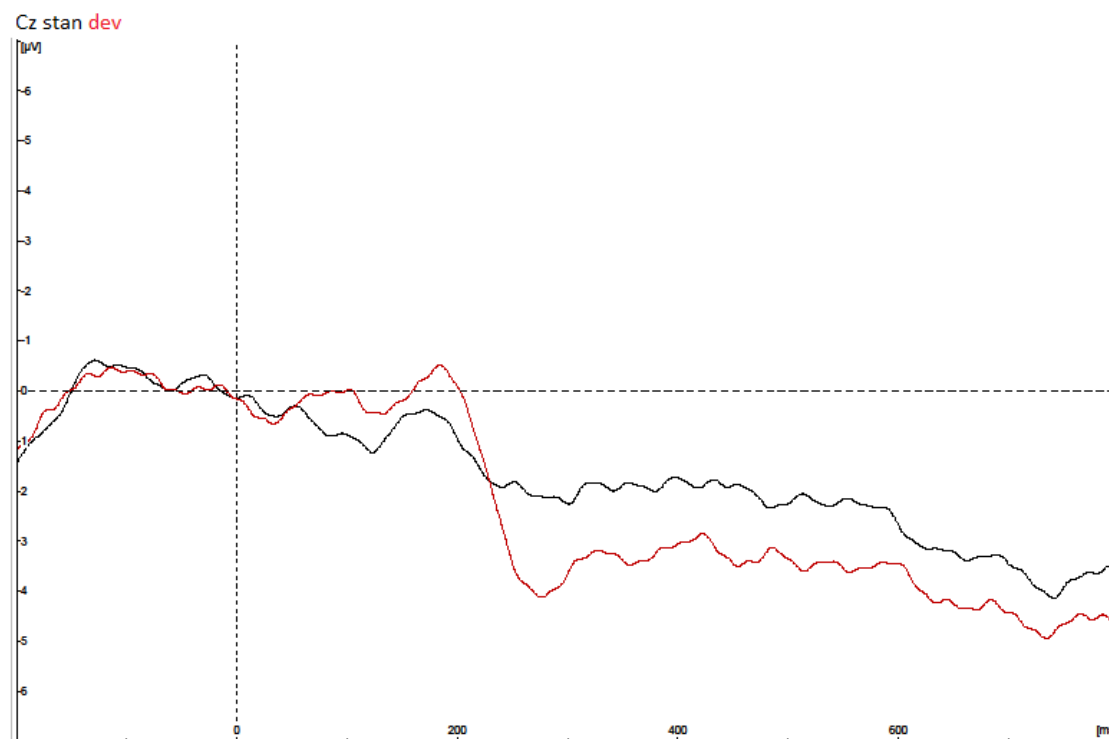

Figure 3b: Grand average (n=18) ERP waveforms at Cz time-locked to the offset of the nucleus of /ba:t/.

100-250 ms interval. Statistical analyses show a main effect of Condition ($F_{1.17}$ = 5,419, p = 0,033, η^2 = 0,242) in the standard 100-250 ms interval for the MMN, which means

that an MMN is evoked. Further we found interactions between the factors Condition and Region ($F_{1.17} = 20,349$, $p = <0,001$, $\eta^2 = 0,545$) and a three-way interaction between Condition, Hemisphere and Region ($F_{1.17} = 6,748$, $p = 0,019$, $\eta^2 = 0,284$). We therefore analyzed each of the quadrants separately: the effect was significant in left and right anterior quadrants ($F_{1.17} = 9,884$, $p = 0,006$, $\eta^2 = 0,368$) and ($F_{1.17} = 12,374$, $p = 0,003$, $\eta^2 = 0,421$) respectively, but not in the posterior quadrants ($p > 0,219$).

50-150 ms interval. Since we observed a double negative peak in our visual analyses we decided to analyze the time-window 50-150 ms, corresponding to the first negative peak, too. Again we found a main effect of Condition ($F_{1.17} = 7,062$, $p = 0,017$, $\eta^2 = 0,293$) with an interaction between Condition and Region ($F_{1.17} = 40,880$, $p = <0,001$, $\eta^2 = 0,706$) and an interaction between Condition, Hemisphere and Region ($F_{1.17} = 14,174$, $p = 0,002$, $\eta^2 = 0,455$). An additional quadrant analysis showed a significant effect in the left and right anterior quadrants ($F_{1.17} = 21,487$, $p = <0,001$, $\eta^2 = 0,558$) and ($F_{1.17} = 18,154$, $p = 0,001$, $\eta^2 = 0,516$), but not in the posterior areas ($p > 0,490$).

120-220 ms interval. The second negative peak that we observed in our visual analyses is in the 120-220 ms time window. In the 120-220 ms interval, statistical analyses also revealed a main effect for the factor Condition ($F_{1.17} = 7,014$, $p = 0,017$, $\eta^2 = 0,292$) as well as a significant two-way interaction between Condition and Region ($F_{1.17} = 22,030$, $p = <0,001$, $\eta^2 = 0,564$), and a three way interaction between Condition, Hemisphere and Region ($F_{1.17} = 5,913$, $p = 0,026$, $\eta^2 = 0,258$). Further post-hoc analysis of the interaction between Condition, Hemisphere and Region in a quadrant analysis revealed a significant effect for frontal regions, both in the left posterior area ($F_{1.17} = 11,092$, $p = 0,004$, $\eta^2 = 0,395$) and in the right anterior area ($F_{1.17} = 16,060$, $p = 0,001$, $\eta^2 = 0,486$).

ERPs time-locked to the onset of /ba:/ stimuli

Based on visual inspection, we observed two responses in the grand average of /ba:/ stimuli at respectively the 200-300 ms and 400-470 ms interval after stimulus onset. The grand average can be found in Figure 4a.
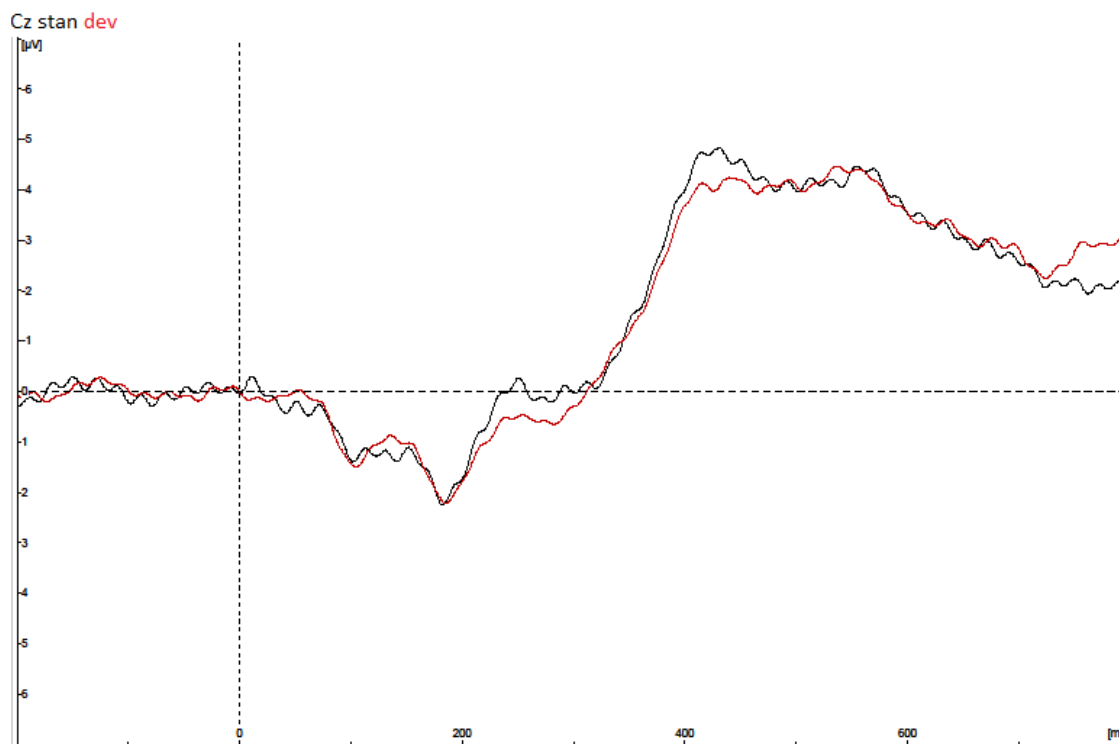

Figure 4a: Grand average (n=18) ERP waveforms at Cz time-locked to the onset of /ba:/.

200-300 ms and 400-470 ms interval. Statistical analysis showed that there are no main effects of Condition in the 200-300 ms interval ($F_{1,17} = 2,987$, $p = 0,102$, $\eta^2 = 0,149$) and 400-470 ms interval ($F_{1.17} = 0,481$, $p = 0,497$, $\eta^2 = 0,028$) or any other interactions (p>0,412). However, it must be noticed that even though these intervals did not reach the level of significance (p<0,05), some participants (n = 12) seem to notice the difference between the two stimuli early in the onset and vowel.

ERPs to /ba:/ stimuli time-locked to the offset of the nucleus

We further examined ERPs time-locked to the offset of /ba:/ stimuli. The grand average reveals a rather unexpected waveform. As can be seen in Figure 4b, the grand average

displays a long positive drift that starts immediately from 0 (=offset nucleus) onwards. We therefore choose the 0-100 ms interval for statistical analysis.
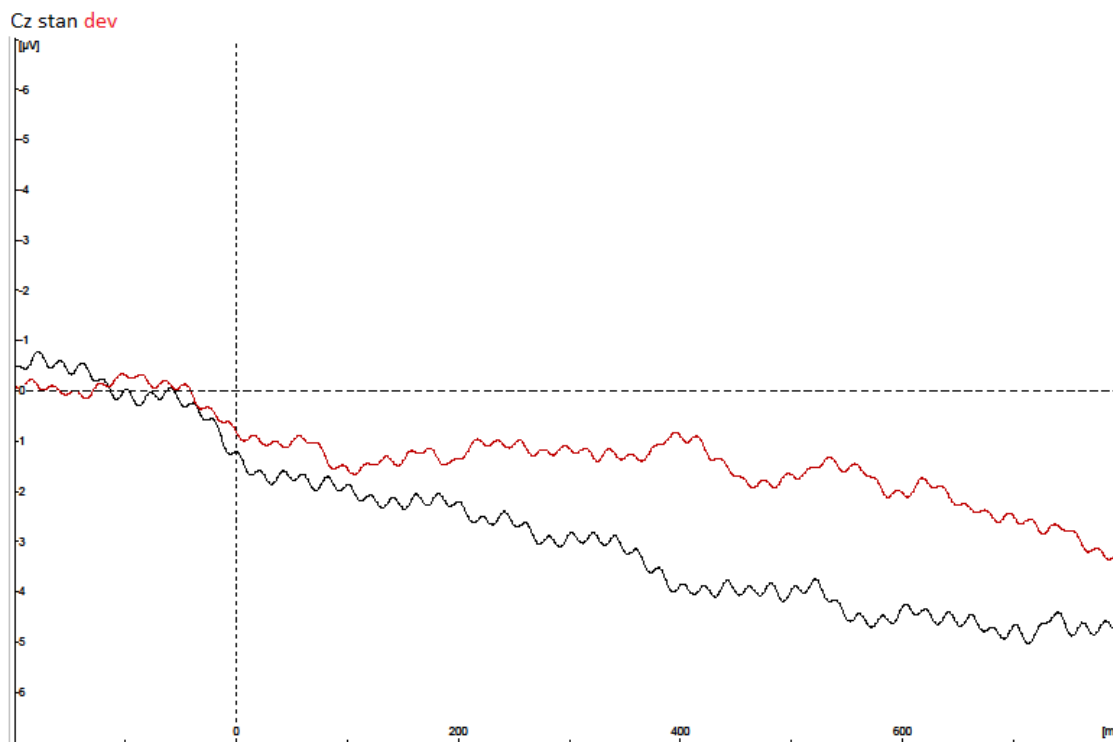

Figure 4b: Grand average (n=18) ERP waveforms at Cz time-locked to the offset of the nucleus of /ba:/.

0-100 ms interval. The time window 0-100 ms turned out indeed to be significant. Statistical analysis shows a main effect of Condition ($F_{1.17} = 6,571$, $p = 0,020$, $\eta^2 = 0,279$), but no significant interactions (p>0,382) were found.

DISCUSSION

In a pilot preceding this EEG study we examined the difference in results between the use of an oddball paradigm with multiple tokens versus one with single tokens. As we expected, the MMN was clearer in an oddball paradigm with one token. However, this does not exclude an MMN to be found in the multiple token variant. Based on our visual analysis we observed in the same time window also an effect in the multiple token condition. From this we can conclude that the memory trace created by acoustic variable standard stimuli is strong enough to evoke an MMN. This is in line with our expectations: in natural speech, the same sound is never pronounced acoustically the same twice. Therefore, acoustic variability is a by-product

of speech production and somehow our brain needs to deal with this and treat these different allophones as variations on the same phoneme of a certain language. The use of multiple tokens in this pilot creates natural speech perception conditions and demonstrates that that our brain indeed treats these variations as the same: a memory trace is created based on acoustic variable standard stimuli and is then violated by a clearly distinct stimulus, i.e. the deviant with coda addition or coda omission, which evokes an MMN.

The main aim of the present EEG study however was to shed more light on the results obtained from the oddball study with 14-month-old infants, done by Nijveld (in prep.). By repeating the same experiment with adult subjects, we tried to clarify and explain their data. In order to be able to compare the results of the present study with the data obtained from the infant study, we strived to create equal conditions for our study. Therefore, we decided to continue with the use of multiple tokens in the EEG study with adults, too. Our expectations were to find an MMN after both coda addition and coda omission for adults, because of a relative low-ranked NOCODA constraint in Dutch Phonology that is not expected to influence the incoming speech any longer.

ERPs time-locked to stimulus onset

In analyzing ERP data time-locked to stimulus onset, we found significant effects when codas were added, but not when they were omitted. Significant effects in the + /t/ condition (e.g. coda addition) were found in the 250-350 ms interval ($F_{1.17} = 8,172$, $p = 0,011$, $\eta^2 = 0,325$) as well as in the 450-550 interval ($F_{1.17} = 7,998$, $p = 0,012$, $\eta^2 = 0,320$). Recall from Table 1 that the mean duration of onset + nucleus (e.g. the onset of the critical time) is 577 ms. Both significant intervals fall within the time window of the vowel, indicating that participants noticed a difference between the standard and the deviant even before the onset of the coda. This means that adults rely on cues provided by the vowel preceding the coda. It remains

unclear why participants perceived these cues only when the coda was added. When the coda was omitted, no significant effects were found in the onset of the stimuli (p>0,102).

ERPs time-locked to the end of the nucleus

Statistical analysis of the present study shows that adults clearly notice an acoustic change from /ba:/ to /ba:t/. As we expected, the ERP of the deviant with coda addition time-locked to the end of the vowel elicited an MMN in the 100-250 ms time window ($F_{1.17} = 5,419$, $p = 0,033$, $\eta^2 = 0,242$). It must be noticed however that the MMN we found was double peaked (cf. figure 3b). This might be a result of the multiple tokens we used. In making separate grand averages of the four tokens, we observed the double peak especially in the second and fourth token of /ba:t/. This is illustrated with the red circles in Figure 5 below. As can be observed from this figure, one early MMN is observed around the 100-200 ms interval and a later effect is observed in the 200-300 ms interval. In the first and third token on the other hand, the effect of the double peak seem to be visible too, but is less clear. In the first token an MMN is observed around 200 ms after. The third token does not evoke an MMN at all.
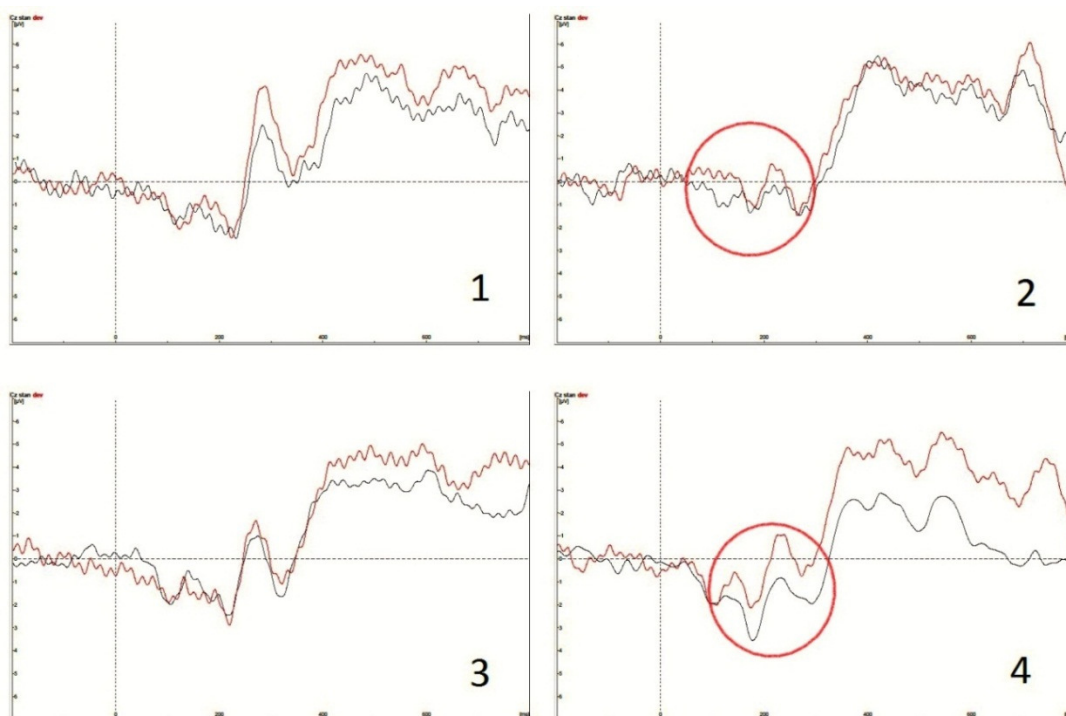

Figure 5: Grand average (n=18) ERP waveforms at Cz of all /ba:t/ tokens separately. The red line corresponds to the deviant, while the black line is the standard sound.

A possible explanation for a double peaked MMN is that one half of the participants shows an MMN in the early time window, while the other half has the MMN in the late time window. The reason that this effect is found only in half of the tokens remains unclear. Further research needs to be done to explain this in more detail.

The results obtained from a change from /ba:t/ to /ba:/ on the other hand, need to be explained further. Instead of the MMN that we expected to see, the ERP to deviant /ba:/ elicited a slow positive drift when the coda was omitted. The same kind of drift has been noticed in the ERP of standard stimuli as well. This positive drift that we observed shows some resemblance with the Closure Positive Shift (CPS) reported by Steinhauer, Alter & Friederici (1999). They observed a large positive deflection with a centro-parietal scalp distribution at intonational phrases and hypothesized that the new found ERP component reflects prosodic processing. In this case, finding such an effect can be regarded as processing expected silence to indicate the end of the stimulus. This is in line with our data: the ERP to standard /ba:/ as well as the ERP to standard /ba:t/ elicits a CPS, which could indicate the expected end of the stimulus. However, finding a CPS instead of an MMN in the ERP to deviant /ba:/ raises some interesting questions on what is going on when the coda is omitted. Note that, in contrast to the ERP to standard /ba:/, the ERP to deviant /ba:/ shows a delayed CPS that starts around 500 ms after the offset of the nucleus (cf. Figure 4b). This delayed CPS reveals that the brain does not assume the deviant /ba:/ to be ended already and is anticipating for coda /t/ in the first 500 ms after nucleus offset. When coda /t/ stays out, a CPS appears just as we observed in the ERP to standard /ba:/.

Comparing the difference between coda addition (MMN) and coda omission (CPS) reveals that the brain responds differently when silence is involved. An interesting sequel to the present study would be to repeat this experiment with fricatives instead of plosives. A plosive consists of multiple phases: first there is silence that corresponds to the airflow that is

stopped in some part of the oral cavity. The airway stays closed, which causes a difference in pressure. Finally, when the closure opens again the released airflow produces a short explosion. Since the coda /t/ that occurred in the preceding standard stimuli also started with a moment of silence it is not so surprising that the brain anticipates for coda /t/ in first instance. The silence of the omission agrees with the onset of the initial expected coda /t/. Fricatives do not consist of such a silence and might therefore elicit a slightly different response. Further research should investigate this in more detail.

Explaining infants' ERP responses to coda addition and coda omission

As we hoped, the data obtained from the current study allows for an explanation of the results of the oddball study with 14-month-old infants, done by Nijveld (in prep.). Results from the infant study show a large resemblance with the results obtained from the present study: the infant's brains shows an MMN when codas are added, but a delayed CPS when codas are omitted. The MMN that is evoked after coda addition demonstrates that infants can detect this acoustic change. The CPS that is found after coda omission on the other hand suggests that 14-month-old infants, just as adults do, are anticipating the coda /t/. In other words, these data illustrate that developing infants do not have difficulty perceiving codas acoustically. The young infant's brain picks up the acoustic changes in stimuli, both for coda addition as for coda omission. Errors in coda perception during a word learn task, as was shown in the study of Levelt (2012), are therefore probably the result of a failure to storing codas correctly into the mental lexicon.

CONCLUSION

Most of the preceding studies to MMN evoked an MMN response in the onset of words. Therefore little attention has been paid to time windows in the offset. In this study, we tried to evoke an MMN in coda position and analyzed both the onset as well as the offset of the stimuli. To summarize, analyzing ERPs time-locked to the onset of stimuli gave us more

insight into when adults notice the difference between the standard and deviant sound. Results showed that most adults noticed the difference between the two types of stimuli during the vowel. This indicates that they detect coda addition early by using cues provided by the vowel.

More importantly, results of the current study highlight the way in which the mature brain deals with coda additions and coda omissions. Statistical analysis ensured that an MMN was obtained when a coda was added to the stimuli, which means that a change from /ba:/ to /ba:t/ is clearly noticed. Interestingly, the brain responds differently when codas are omitted and a silence therefore replaces the coda. Instead of the MMN that we expected, a delayed CPS is evoked to deviant /ba:/, which indicates that the brain is anticipating to coda /t/. Although responses to coda addition and omission manifest themselves in different ways, we may conclude that adults notice both changes. Since adults do not have problems with coda perception and noticed the change in both directions, these data confirm our hypotheses that NOCODA constraints do not influence the incoming speech for adults any longer.

In addition, results of the pilots preceding this study demonstrated that there is no difference in results using an oddball paradigm with single tokens versus one with multiple tokens. Therefore we could continue the experiment using an oddball paradigm with multiple tokens, just as was done in the oddball study with 14-month-old infants. Doing so allowed us to compare both studies. Comparison with the data obtained from adult subjects showed that the infant's brain responds in the same way to coda additions and coda omissions as the mature brain does (an MMN when codas are added and a CPS when the codas are omitted). This means that infants note changes in both directions and that insensitivity to coda omissions are probably the result of a failure to store codas correctly in the mental lexicon.

REFERENCES

Boersma, P., & Weenink, D. (2005, 2012) Praat: doing phonetics by computer (Version 4,3

and 5.3.04) [Computer program]. Retrieved from http://www.praat.org/.

Böttcher-Gandor, C., & Ullsperger, P. (1992). Mismatch negativity in event-related potentials

to auditory stimuli as a function of varying interstimulus interval. *Psychophysiology*,

29, 546–50.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English.* Harper & Row, New York.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P.

Shah (Eds.), *Models of working memory: mechanisms of active maintenance and

executive control*. Cambridge: Cambridge University Press.

Eulitz, E., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological

representations in the mental lexicon during speech recognition. *Journal of cognitive

neuroscience*, 16 (4), 577-83

Fischer, C., Morlet, D., Bouchet,P., Luaute, J., Jourdan, C., & Salord, F. (1999). Mismatch

negativity and late auditory evoked potentials in comatose patients. *Clinical

Neurophysiology*, 110 (9), 1601-1610.

Fitzpatrick, J., & Wheeldon, L. R. (2002). Phonology and phonetics in psycholinguistic

models of speech perception. In N. Buton-Roberts, P. Carr, & G. Docherty (Eds.),

Phonological knowledge (pp. 131–159). Oxford: Oxford University Press.

Garrido, M.I., Friston, K.J., Kiebel, S.J., Stephan, K.E., Baldeweg, T., & Kilner, J.M. (2008).

The functional anatomy of the MMN: a DCM study of the roving paradigm.

Neuroimagine, 42, 936–44.

Garrido, M.I., Kilner, J. M., Stephan, K. E. & Friston, K. J. (2009). The mismatch negativity:

a review of underlying mechanism. *Clinical Neurophysiology*, 120, 453–463.

Jääskeläinen, I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levänen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci USA*, 101, 6809–6814.

Levelt, C.C. (2012). Perception mirrors production in 14- and 18-month-olds: the case of coda consonants. *Cognition*, 123, 174-179.

Levelt, C.C., Schiller, N.O., & Levelt, W.J.M. (2000). The acquisition of syllable types*. Language Acquisition*, 8 (3), 237-264.

Menn, L. (1978). Phonological units in beginning speech. In A. Bell & J. Hooper (Eds.), *Syllables and segments* (pp. 157–172). Amsterdam: North-Holland.

Näätänen, R. (1992). *Attention and Brain Function*. Erlbaum, Hillsdale.

Näätänen, R. (2001). The Perception of Speech Sounds by the Human Brain as Reflected by the Mismatch Negativity (MMN) and its Magnetic Equivalent (MMNm). *Psychophysiology*, 38, 1-12.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42, 313–329.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118, 2544-2590

Näätänen, R. & Winkler, I. (1999) The concept of auditory stimulus representation in neuroscience. *Psychological Bulletin*, 125, 826–59.

Nijveld, A. D. (in prep.) *Coda perception in 14-month-old infants: ERP evidence*. ResMA Thesis, Leiden University, Leiden.

Prince, A., & Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers University Center for Cognitive Science Technical Report 2.

Smith, N. (1973). *The acquisition of phonology: A case study*. London: Cambridge University

    Press.

Steinhauer, K., Alter, K., & Friederici, A.D. (1999). Brain potentials indicate immediate use

    of prosodic cues in natural speech processing**.** *Nature Neuroscience,* 2 (2), 191-196

Werker, J., Cohen, L., Lloyd, V. & Casasola, M. (1998). Acquisition of word-object

    associations by 14-month-old infants. *Developmental Psychology, 34, 6*, 1289-1309.

APPENDIX: pilot studies

Prior to the current experiment we did two pilot studies in order compare results of an oddball paradigm with single tokens versus one with multiple tokens. Details about participants, stimuli and the results of visual analyses are described in the sections below. The experimental design and procedure were the same as for the study described in this paper and are therefore not described in the appendix. For further details about the experimental design and procedure, please read sections 2.3 and 2.4.

Participants

A total of nineteen adults participated in the pilot studies, of which six participants were later excluded from analysis due to technical problems (n=2) or because of insufficient trials that could be averaged in one of the conditions (less than 70 trials, n=4). The remaining thirteen participants were between 19 and 30 years old (mean age = 23,34, women = 6) and participated either in the single token condition (n= 6, women = 3) or in the multiple token condition (n=7, women = 3). All subjects were native speakers of Dutch and were righthanded. None of them reported hearing or neurological disorders in their history. Participants gave informed consent prior to the experiment and were paid for their participation or received university credit points.

Stimuli single token condition

For the one token condition we used one /ba:t/ token that we had not used in the multiple token condition, but was recorded by the same speaker in the same session. This token had a duration of 839 ms, with the end of the vowel at 490 ms, and the onset of the coda starting at 596. (Note that this is within the range of stimuli used for the multiple token condition). Using a a speech waveform editor (*Praat*; Boersma & Weenink, 2005), we subsequently manipulated this token by splicing the onset with vowel from the coda. To make this sound more natural, we then faded out the last 50 ms from the nucleus, which was not discernible.

This manipulated sound was judged to sound natural by two independent Dutch monolingual speakers. See Figure 6 below, with on top the original /ba:t/-token, and below the spliced /ba:/-token. Acoustic details of these stimuli are presented in Table 3 below.
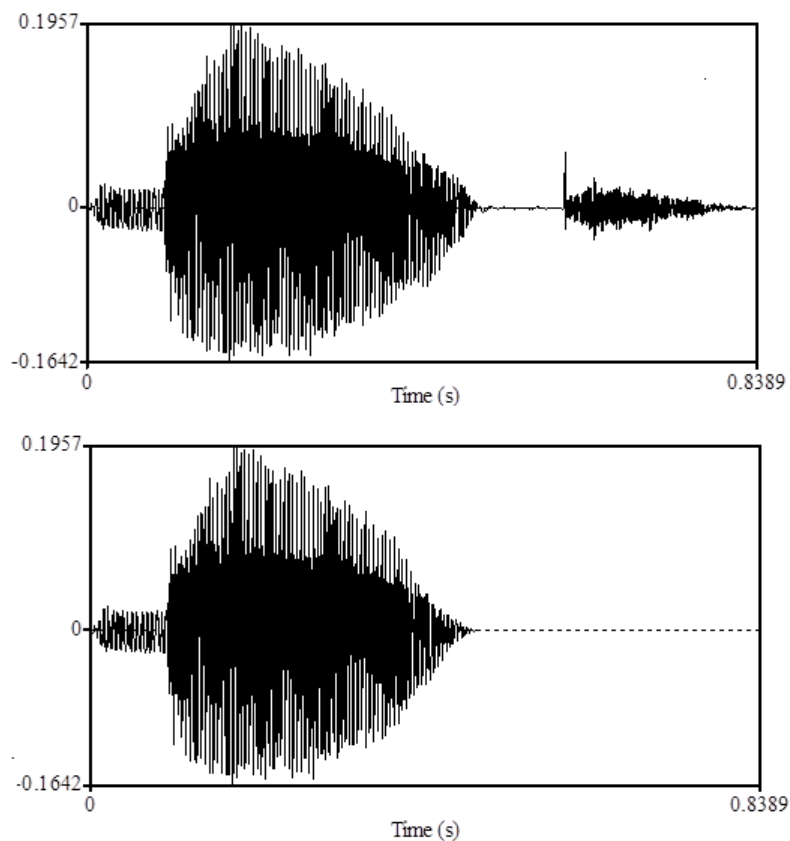

Figure 6: oscillograms of the stimuli used in the single token condition.

	Duration /b/	Duration vowel	Onset ct. (/ba:/)	Duration silence	Onset /t/	Duration /t/	Total duration
/ba:t/	109	436	535	62	596	242	839
/ba:/	109	381	-	-	-	-	490

Table 3: acoustic details of single token stimuli in ms.

Stimuli multiple token condition

The stimuli of the multiple token condition consisted of four tokens of the pseudowords /ba:/ and /ba:t/. They were the same as those that are described in section 2.2. Please see Table 1 in section 2.2 for acoustic analysis on these stimuli.

Results

We did not carry out statistical analyses. Results of the pilot studies are based on visual

analysis. Grand averages of standard and deviant stimuli at Cz are displayed in Figure 7

(single token condition) and Figure 8 (multiple token condition).
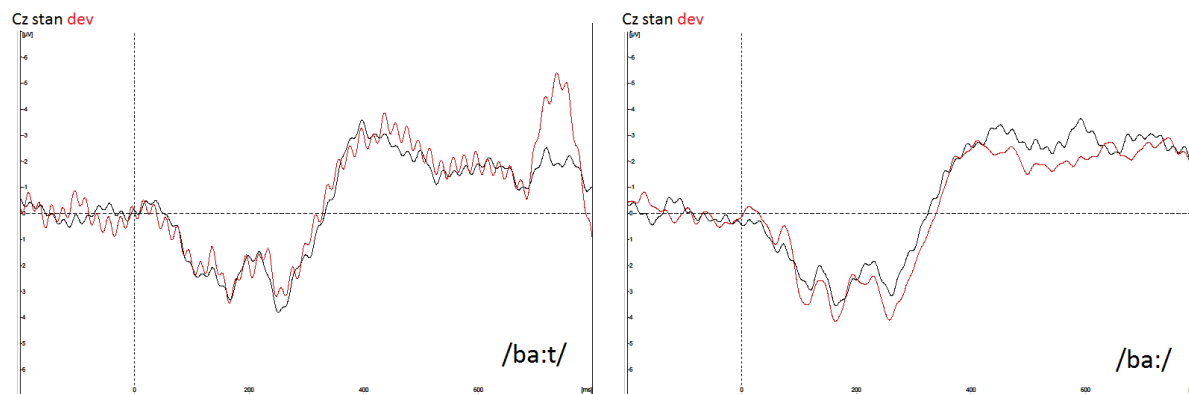

Figure 7: Grand average (n=6) ERP waveforms at Cz of the pilot study with single tokens.
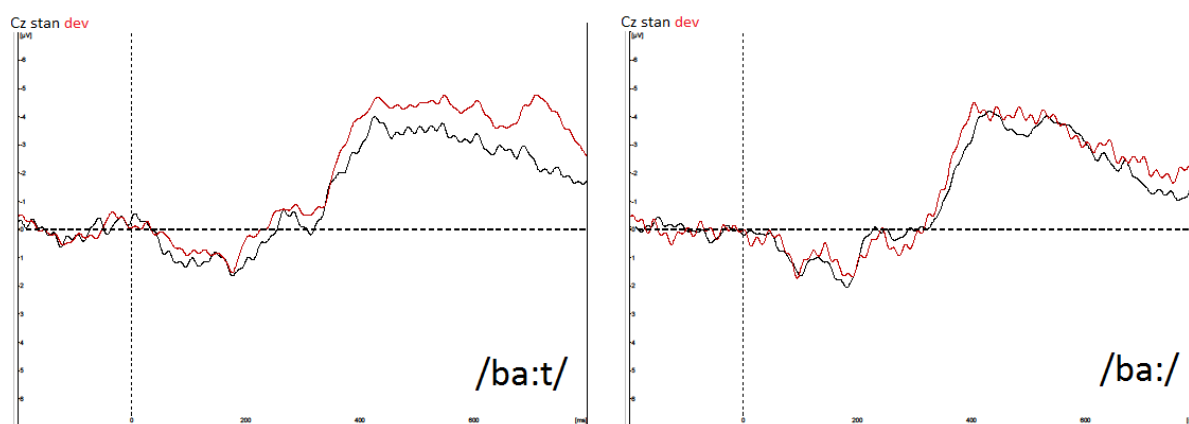

Figure 8: Grand average (n=7) ERP waveforms at Cz of the pilot study with multiple tokens.

As can be seen from Figure 7, the ERP to deviant /ba:t/ elicits a clear response around 700 ms

after stimulus onset. To put it differently, the response is elicited roughly 165 ms after the

onset of the critical time, i.e. the end of the vowel, (cf. Table 3). This falls exactly within the

time window in which we expect an MMN to occur (100-250 ms after the onset of the

deviant stimulus). A similar, but less clear, negative peak can be observed for deviant /ba:t/ in

the multiple token condition (Figure 8): an MMN is evoked in the 600-700 ms time window.

Note however that no clear effect is found in either paradigm when the coda is omitted. More

subjects are needed to better understand how the brain reacts to coda omissions. Continuation with one of the oddball paradigms in the following experiment will hopefully clarify this.

In conclusion, results show that an MMN is evoked in both conditions. This means that acoustic variability does not prevent a memory trace to be built. We chose to continue the study with multiple tokens, in order to be able to compare the results to the oddball study with 14-month-old infants that used multiple tokens, too.