Looking at the *knife* when hearing “map”

How important is the Dutch mental lexicon when processing English words in noise?

Name:	Maria van de Groep
Student number:	s4363892
Date of completion:	13 March 2017
Type of report:	Bachelor thesis
Supervisor:	Dr. Florian Hintz
2nd reader:	Dr. Sharon Unsworth

Index

0. Abstract

When recognizing spoken words, similar sounding words are temporarily considered target candidates. Previous research has shown that words overlapping with a spoken target in phonological onset compete for recognition for as long as the acoustic overlap lasts. Moreover, it has been shown that such phonological competition increases when recognizing words in a foreign language, due to native and non-native words being activated simultaneously. Also, relative to perceiving words in the clear, phonological competition appears to be increased when words are processed in noise. This most likely reflects that listeners interpret the speech signal with more flexibility, maintaining target competitors active for a longer period of time.

The present study aimed to investigate the relative importance of the native and non-native mental lexicon when recognizing non-native speech in noise. Eye movements of native Dutch participants were recorded as they listened to English words while looking at a display containing four objects. On filler items, the visual referent depicting the spoken target was present, along with three unrelated distractors. On experimental items, the picture of the spoken target word was absent. Instead, the display featured an English (but not Dutch) phonological onset competitor, a Dutch (but not English) phonological onset competitor and two unrelated distractors. After a preview phase of three seconds, the spoken target was presented via headphones and participants were instructed to indicate whether or not it was visually depicted. On half of the items the spoken word was masked by speech-shaped noise; on the other half, the word was presented in the clear. Participants' looks to the English and Dutch competitors on experimental items were analyzed, starting at the onset of the spoken target.

Unlike previous studies, I did not observe increased fixations to either Dutch or English competitors during early time windows (200-600 ms after word onset). However, significant biases did appear in later time windows (900-1500 ms). In the clear condition, a significant bias for the English onset competitor was found between 900 and 999 ms. A bias for the Dutch onset competitor was found between 1400 and 1499 ms. In the noise condition, I found a bias for the Dutch onset competitor during the same time window as in the clear condition. The bias for the English onset competitor appeared 100 ms later than in the clear condition, and was only significant by subjects (not by items). These results imply that the importance of the non-native mental lexicon decreases once the spoken input is masked by noise. The importance of the native Dutch mental lexicon when listening in noise is comparable to listening in the clear.

1. Introduction

In the process of understanding speech, it may on some occasions seem more difficult and feel as if it takes more of an effort to correctly comprehend what your interlocutor is saying. Even under ideal circumstances, processing speech is a theoretically complex process: as soon as a word starts to unfold, multiple words with the same phonological onset become partially active as possible target candidates. The more sounds you hear, the more possible candidates are eliminated, until the target word is left. This process takes place automatically, and does not take much of an effort. But when listening to speech in a noisy environment – like a train station or a cafe – it seems to be harder to correctly understand what is being said than in an environment without any background noise. Another example of speech processing in adverse conditions is when you are having a conversation with an interlocutor who is speaking in a language which is not your native tongue, even if you are able to speak this language fluently. The present study looks at the interaction of these two factors, by investigating foreign language comprehension in noise. My specific research question was: What is the importance of the mental lexicons of both the native and the non-native language when listening to non-native speech in clear and in noise? To tackle this research question, an eye-tracking experiment was conducted with native Dutch participants who listened to English words either occurring in the clear or masked by speech-shaped background noise.

First, the literature on spoken-word recognition is reviewed. Next, the *visual world paradigm*, the experimental paradigm used in this study, is explained. Then, previous research into both non-native speech processing and speech processing in noise is presented. After formulating the hypotheses, the methods and results are described and discussed. Lastly, a conclusion is drawn, along with suggestions for further research.

1.1 Spoken-word recognition

The recognition of words in a continuous stream of speech is a theoretically complex process. Many researchers have attempted to explain how this process works by developing models for spoken-word recognition. Although these models are mostly of a theoretical nature, some of them have been converted into computer-based mathematical models in order to test their validity.

One of the first models for spoken-word recognition was the Logogen model by Morton (1969). In this model, each word is represented by a *logogen*. “The logogen is a device which accepts information from the sensory analysis mechanisms concerning the properties of linguistic stimuli and from context-producing mechanisms.” (Morton, 1969:165) The model takes spoken input (and context information) as a base, and uses this to activate word form representations that have been stored in a certain logogen before. Processing units gather information from auditory input and fire when a threshold is reached. Only when the threshold is reached, the word meaning becomes accessible. Thresholds can differ per logogen, and are usually lower for high frequency words. Subsequent research has shown that this model is incomplete as it does not presuppose any connections between word form representations.

The TRACE model for spoken-word recognition (McClelland & Elman, 1986) does not only take the bottom-up acoustic signal as its input, but also works with top-down activation. Acoustic features, phonemes and words are each represented by a processing layer, with separate units for each feature, phoneme or word. As certain units become activated by auditory input, they activate units in higher layers of processing. These higher units can then also give top-down feedback to lower layer units. In the TRACE model, activation of processing units is cascaded (as opposed to the threshold activation in the Logogen model); as soon as there is input, units also begin to send output to other processing units. Apart from top-down activation, the TRACE model also differs from other models in that it makes use of *lateral inhibition*; because acoustic features can only be part of one unit at a time, the activation of one unit will cause that unit to decrease the activation of other units within one processing layer. This causes the units within one layer to compete for activation until one unit has received the most auditory input and is selected.

The COHORT model (Marslen-Wilson, 1987) assumes three processes to take place sequentially during spoken-word recognition. During the first process, *access*, several different word forms are activated based on the spoken input, and by spoken input only. In this process, context doesn’t play a role. The first group of selected words is called a *cohort*. During the process of *selection*, context information is also regarded: the words in the cohort are sorted by quality of the match with both auditory input and context. During the process of *integration*, the features of the selected words are integrated into the entire utterance with regard to factors such as grammatical class and semantics. As TRACE, the COHORT model assumes spoken-word recognition to be cascaded: the auditory input and the words in the cohort are matched constantly in processing cycles, until the point where only one of the selected words is left. This point is called the *recognition point*, which can be reached before the spoken word has fully unfolded.

The Distributed COHORT model (DCM) (Gaskell & Marslen-Wilson, 1997) is an adjustment to the COHORT model described above. The DCM takes phonetic features as its input, which go through both processing and context units. The context units store copies of the activation in between processing cycles. In this way, the most recent activation is updated constantly. Next, the current activation pattern is run through two additional processing units: one that represents phonological word forms and one that represents word meanings.

The models for spoken-word recognition described above are just a few of the many models which have been developed over the years. However, they do give a general idea of how the process of spoken-word recognition works. An important aspect is that word recognition happens in cycles. It is a constant process where information is evaluated and re-evaluated until the point where one word is selected as the best possible option: *the recognition point*. Another important aspect is that several words compete for activation before one word is selected.

1.2 The visual world paradigm

In most models of spoken-word recognition, the selection of the first group of possible target candidates is based on a shared phonological onset. That is, when a spoken word is being heard, but has not unfolded completely, the auditory input available is that of the first phonemes. Based on these first phonemes, several possible target candidates are selected. As the phonemes of a word unfold, the number of possible candidates decreases. Although the nature of this assumption is theoretical, several studies have given strong evidence for *phonological onset competition* (e.g., Allopenna et al., 1998; Dahan & Gaskell, 2007; Huettig & McQueen, 2007). One way of investigating phonological onset competition is by using the *visual world paradigm*.

The *visual world paradigm* is a theoretical assumption which relies on the relationship between eye movements and speech processing. As speech is processed, eye movements are assumed to reflect the integration of visual information with auditory information (with a slight delay of 200 ms between spoken input and the movement of the eyes). Measuring eye movements with the use of an eye tracker therefore is a widely used technique for research concerning spoken-word recognition. Eye-tracking experiments generally consist of a display containing four pictures and a spoken target word. This spoken target word can be either visible on the display as a picture, or not, depending on the experimental design. If present, the target picture is typically accompanied by one or more competitors, as well as one or more unrelated distractors. For example, competitors could overlap with the spoken target in phonological onset. Participants can be instructed to simply look at the display and listen to the spoken input, but another possibility is to give them a specific task, for example, to click on the picture of the word they just heard. Researchers have also manipulated the spoken input. The spoken target word can be presented in isolation, or it can be embedded in carrier sentences. The analysis of eye-tracking data usually consists of calculating the proportion of fixations to the target picture and to the competitor(s), relative to fixations to the distractor pictures at multiple moments in time. The outcome is the probability of fixations to either the target or competitor pictures over time. In this way, eye-tracking experiments make the incremental processing of speech visible.

A seminal study using visual world eye-tracking was done by Allopenna, Magnuson and Tanenhaus (1998), who investigated on-line word processing, specifically phonological onset competition. Participants in this study were presented with a grid on a computer screen with four shapes on it (a square, a diamond, a circle and a triangle). The grid also contained four objects: a target object (e.g., *beaker*), a phonological onset competitor (e.g., *beetle*), a rhyme competitor (e.g., *speaker*) and an unrelated distractor (e.g., *carriage*). The participants were instructed via headphones to “Pick up the beaker. Now put it on the diamond.” As participants heard these instructions, their eye movements were measured. Measured from the onset of the spoken target word, pictures of the phonological onset competitor and the target word were looked at for as long as the acoustic overlap lasted (about 300 ms). This most likely reflected that the competitor and the target word both competed for activation during this time period.

1.3 Non-native spoken-word recognition

A great deal of previous research into spoken-word recognition using the *visual world paradigm* has focused on the processing of non-native speech. The interference of the non-native lexicon when processing native speech by bilinguals starts as early as age two (Bobb et al., 2016). Bobb and colleagues investigated the role of the native and non-native lexicon in bilingual toddlers (aged 21 to 61 months), using a visual world set-up as described above. They found that bilingual toddlers looked at competitors from the non-native language when processing native speech more than their monolingual peers did. Although the present study is not focused on bilinguals who acquired languages at an early age, but rather on people who acquired a second language at a later age, the study by Bobb and colleagues shows that interference from the non-native lexicon can impact language processing at an early stage.

Weber and Cutler (2004) conducted a study concerning lexical competition in non-native word processing. With an eye-tracking experiment they confirmed the presence of between-language competition. They presented both Dutch and English natives with a grid on a screen with four shapes and four pictures on it, very much like the study by Allopenna and colleagues (1998). This grid included a picture of a target word (e.g., *desk*) and a picture whose Dutch (but not English) name was a phonological onset competitor (e.g., *deksel*, English: *lid*). Dutch natives fixated both the target picture and the Dutch competitor more than they fixated the distractors. The Dutch native participants experienced increased competition as compared to the English natives, who did not show a bias for the Dutch competitor *deksel* (Weber & Cutler, 2004). These results imply that processing non-native speech leads to increased phonological competition in comparison with the processing of native speech. This may be due to the fact that, when processing non-native speech, not only phonological onset competitors that originate in the non-native mental lexicon compete for activation, but also those from the native mental lexicon.

In a similar vein, Marian and Spivey (2003) conducted two experiments testing native speakers of Russian who had English as a second language. In the first experiment they tried to ensure an English language environment: all communication with participants was done in English, and no attention was paid to the native language of the participants. Participants were presented auditorily with an English target word, which was also depicted on the screen. The screen also contained pictures of both an English and a Russian phonological onset competitor. They found that participants biased both English and Russian competitors over distractor pictures. When processing a non-native language, words from both the native and non-native language competed for activation. In the second experiment Marian and Spivey (2003) tried to achieve a Russian language environment, which was the participants' native language: all communication was done in Russian. Participants heard a Russian target word, which was also depicted on the screen. The screen displayed pictures of Russian and English phonological onset competitors. In the second experiment, Marian and Spivey found biases for the Russian competitors, but not for the English competitors. They argued that, when processing a native language, no competition from non-native competitors arises. The interference of one language when processing the other language seems to be stronger in non-native listening than in native listening.

However, there have been studies that showed effects of non-native phonological competition when listening in the native language. Lagrou, Hartsuiker and Duyk (2013) conducted two experiments with native Dutch participants who had started learning English at high school age. In the first experiment, participants were presented with non-native English carrier sentences that contained an English target word (e.g., *flower*). This target word was depicted on a screen, as well as a Dutch phonological onset competitor (e.g., *fles*; English: *bottle*), and two unrelated distractors. Both the target word and the competitor could be inserted into the carrier sentence and were not predictable. Participants were given a look-and-listen task. The first experiment confirmed that words from both the native and the non-native mental lexicons compete for activation in non-native speech processing. In the second experiment in this study, the authors reversed the target language and the inter-language competitors: Dutch participants were presented with Dutch carrier sentences and target words, and (non-native) English onset competitors. They found that participants looked more at the English competitors than at the distractors, but only when the phonological overlap between competitor and target word was large enough.

Together, these studies indicate that processing non-native speech – with speech being either single words or words embedded in sentences – modulates the processes underlying spoken-word recognition. When recognizing non-native speech, both the native and non-native lexicons compete for activation. The activation of these additional target candidates results in increased lexical competition for non-native listeners, when compared to native listeners.

1.4 Spoken-word recognition in noise

Another factor that impacts the processes underlying spoken-word recognition is the presence of background noise. A key study showing the effects of noise on speech processing is the one by Brouwer and Bradlow (2016). They replicated the study by Allopenna and colleagues (1998). The authors presented participants with a grid featuring four shapes and four pictures: one was the target (e.g. *beaker*), there was one phonological onset competitor (e.g. *beetle*), one rhyme competitor (e.g. *speaker*) and one unrelated distractor (e.g. *carriage*). Participants' eye movements were then measured as they heard the instructions: "Pick up the beaker. Now put it on the diamond." Unlike the study from Allopenna and colleagues (1998), this was done in a clear *and* a noise condition. Relative to the clear, both the phonological onset competitor and the rhyme competitor competed for activation for a longer period of time when the target word was presented in noise. This may reflect that participants process the auditory input more flexibly when background noise is present, keeping competitors active for a longer period of time. Similarly, Ben-David and colleagues (2011) replicated the study by Allopenna et al. (1998) in younger and older adults, mixing the sentence recordings with speech-shaped background noise. As Brouwer and Bradlow, they observed that both younger and older adults were less accurate in recognizing the spoken target word when it was presented in noise. Moreover, relative to presentation of the same word in clear, the presence of background noise resulted in elongated phonological onset competition.

1.5 Hypotheses

All in all, as reviewed above, processing non-native speech and processing speech in noise affect the competition dynamics underlying spoken-word recognition. The question I asked in this study concerns the interactive effects of both adverse conditions. In other words: how strong is the influence of the native and non-native lexicon when processing non-native speech in noise? To investigate this question, Dutch participants took part in an eye-tracking experiment and were presented with English words, while they looked at a screen with four objects. On experimental trials, the target word (e.g. "map"; Dutch: kaart) was presented auditorily, and was not visible as a picture on the screen. One of the objects on the screen overlapped with the English target word in Dutch (but not English) phonological onset; this was the Dutch competitor (e.g. *mes*; English: *knife*). Another object on the screen was the English competitor (e.g. *match*; Dutch: *lucifer*), which overlapped with the target word in the first syllable of the English name. The other two objects were completely unrelated distractor items (e.g. *turtle* and *bed*). On filler items, a depiction of the spoken target was presented next to three unrelated distractors. On half of the items, the target word was masked by speech-shaped noise, the other half was presented in clear. Participants' task was to listen to the spoken target carefully and indicate whether or not it was among the displayed objects.

In line with previous research, measuring from the onset of the spoken target word, I hypothesized that participants would look at both the Dutch and English competitors in the clear condition, (Weber & Cutler, 2004; Marian & Spivey, 2003; Lagrou et al., 2013). For the noise condition, there were multiple possibilities. Participants could be focusing more on the English competitor, thereby neglecting the Dutch onset competitor, since English was the language of the experiment. Another possibility was that, relative to the clear condition, participants would 'fall back on their native language when listening in noise', reflecting an attempt to exploit the greater experience they have in native listening. Lastly, looks to both the English and Dutch competitors could increase, which would reflect that the effects of non-native speech processing and speech processing in noise are additive and that participants keep possible target words activated for a longer period of time.

2. Method

2.1 Participants

Twenty participants took part in the experiment. All participants were native speakers of Dutch with high proficiency in English. All participants started to learn English in high school, aged twelve. A standardized language proficiency test (LexTALE, Lemhöfer & Broersma, 2012) was used to assess each participant's English proficiency level. Participants' average percentage score was 78.1 (*SD* = 13.2, see Lemhöfer & Broersma, 2012, for discussion). None of the participants reported a history of either developmental or acquired speech, hearing or brain problems. All participants had normal or corrected-to-normal vision.

2.2 Design

Using an eye-tracker, participants' eye movements were measured as they listened to an English word while looking at four pictures on the computer screen. On experimental trials, participants heard a target word and saw a picture of an object overlapping in its English name with the spoken English target word in phonological onset. In the same display they saw a picture, whose Dutch (but not English) name overlapped with the spoken English target in phonological onset. The remaining two pictures were distractors, unrelated to the spoken target and the competitor objects. On filler trials, the participants heard an English target word and saw a picture of that word, together with three unrelated distractors. The dependent variable in the experiment was the probability of fixations to each of the objects present on the screen, as measured from the onset of the spoken target. Trials were either presented in the clear or masked by speech-shaped background noise.

2.3 Materials

Twenty-two experimental items were composed (see Appendix), each of which consisted of five words. One word within an item, the target word, was presented auditorily (e.g., "map"). The remaining four words were presented as picture objects. Pictures were either retrieved from the BOSS database (Bank of Standardized Stimuli, Brodeur et al., 2010) or from the Google search engine. One of these objects was the English phonological onset competitor whose English, but not Dutch name overlapped with the target in phonological onset (e.g., English: *match,* Dutch: *lucifer*). Another object was the Dutch phonological onset competitor. The Dutch but not the English name of that object overlapped with the English target in phonological onset (e.g., Dutch: *mes* English: *knife*).

The English target and the English onset competitor overlapped on average in 2.18 phonemes (*SD* = 0.395); the English target and the Dutch competitor overlapped on average in

2.18 phonemes (*SD* = 0.501). A paired t-test revealed no differences in the amount of phonological overlap between the target and both competitor types ($t(21) = 0.00$, $p = 1.00$). Additionally, the English and Dutch onset competitors did not differ with regard to the number of phonological neighbors ($t(21) = 1.740$, $p = .096$) or the mean frequency of these neighbors ($t(21) = 1.437$, $p = .165$) as determined using the ClearPond database (Marian et al., 2012).

All five words within one item had the same number of syllables (9 items with one-syllable words; 13 items with two-syllable words) and were comparable in word frequency (mean target = 38.61, mean English competitor = 27.80, mean Dutch competitor = 10.34, mean distractors = 79.80; $F(4, 103) = 3.576$, $p = .009$). Note that the statistical effect of word frequency was driven by the Dutch competitors featuring a lower frequency than all the other words. However, this did not undermine our hypotheses as I predicted that a bias towards the English competitors would be larger than a bias towards the Dutch competitors, since the spoken language is English. Any bias towards (low frequency) Dutch competitors would thus provide even stronger evidence for between-language competition. Each item was hand-checked for semantic and visual similarity.

Twenty-two filler items were used. Each consisted of four English words: one target word, which was presented both auditorily and on the screen, and three completely unrelated distractors. The English or Dutch name of the target did not overlap with any of the distractors phonologically (in either English or Dutch) or at any other level.

For the recording of the spoken input, I asked a native speaker of British English to read all of the 44 target words in a neutral intonation. Recordings and audio editing was done using Audacity. A second version of each file was created by adding speech-shaped noise, with a signal-to-noise ratio (SNR) of +3 dB to the spoken word.

2.4 Procedure

Participants were asked to fill out a short questionnaire assessing their language background and their auditory and visual abilities. The experiment was administered in a sound-damped booth. The eye-tracker was calibrated. The trial structure was as follows: Participants were presented with a fixation dot (to ensure that they fixated the center of the screen), followed by a preview phase of 3000 ms for the participants to look at the four pictures on the screen. After that, the spoken word was played back via headphones and the participants were asked to indicate whether or not the target object was among the four objects by pushing either the left (not present) or right (present) CTRL-button on the keyboard.

Experimental and filler items were either presented in the clear or masked by speech-shaped noise. The two versions of one item were distributed over two experimental lists, with an equal number of experimental and filler items and an equal number of clear and noise trials on each list. Clear and noise trials were presented in blocks. The order of the blocks was randomized for each participant. Each participant was randomly assigned one list. Each participant saw all of the 44 trials on one list.

2.5 Data analysis

Items for which an incorrect or no response had been given were removed. Participants' eye movements were analyzed starting at the onset of the spoken target onset, until 1600 ms post onset. The data were binned into 100 ms time windows. For each time bin, I calculated ratios by dividing fixations to the critical objects (e.g., target, English competitor, Dutch competitor) by the fixations to the average of the distractors. These ratios were log-transformed using R (R Core Team, 2015). A ratio of zero means equal looks to the critical objects and averaged distractors; a value larger than zero implies a bias for the critical objects. I subsequently carried out paired t-tests between the 0-99 ms (baseline) time bin and all subsequent time bins. As it takes around 200 ms to program and launch a saccadic eye movement (Saslow, 1967), we can assume that eye movements during the 0-99 ms time bin were not influenced by processing of the spoken targets. A difference in fixation behavior between the baseline and the subsequent time bins can thus be assumed to reflect the cognitive processes underlying spoken-word recognition (in noise).

3. Results

Out of a total of 880 trials, 84 trials had to be removed as participants gave an incorrect (68) response or no response at all (16). Independent samples t-tests showed that there was no significant difference with regard to the number of excluded trials in the various conditions (clear vs. noise: experimental trials (t_1 (38) = 0.146, p = .885; t_2 (42) = 0.109, p = .914); filler trials (t_1 (38) = -0.470, p = .641; t_2 (42) = -0.365, p = .717)).

Figures 1 and 2 present participants' eye movements on filler trials. The graphs show the probabilities of fixations to the target word and to the averaged distractors. In the clear condition (Figure 1), a bias for the target word arose at around 600 ms after word onset (t_1 (19) = -1.925, p = .035, one-tailed; t_2 (21) = -2.161, p = .042). In the noise condition (Figure 2) the same target objects were biased slightly later, at around 800 ms after word onset (t_1 (19) = -2.379, p = .028; t_2 (21) = -2.370, p = .027).

Figures 3 and 4 plot the results for the experimental trials, displaying probabilities of fixations to the English phonological onset competitor, the Dutch phonological onset competitor, and the average of the two distractors. In the clear condition (Figure 3), participants biased the English phonological onset competitor over the unrelated distractors between 900 ms and 999 ms after word onset (t_1 (19) = -2.081, p = .026, one-tailed; t_2 (21) = -1.801, p = .043, one-tailed). A bias for the Dutch phonological onset competitor was observed between 1400 and 1499 ms (t_1 (19) = -1.941, p = .034, one-tailed; t_2 (21) = -2.300, p = .032).

In the noise condition (Figure 4) there is a similar pattern in the biases for the phonological onset competitors. Participants biased the English phonological onset competitor over the unrelated distractors between 1000 and 1099 ms. This difference appeared slightly later than in the clear condition, and was only significant by subjects; not by items (t_1 (19) = -1.915, p = .036, one-tailed; t_2 (21) = -1.168, p = n.s.). The Dutch phonological onset competitor showed a bias in the same time window as in the clear condition (1400-1499 ms: t_1 (19) = -2.474, p = .023; t_2 (21) = -1.967, p = .032, one-tailed).

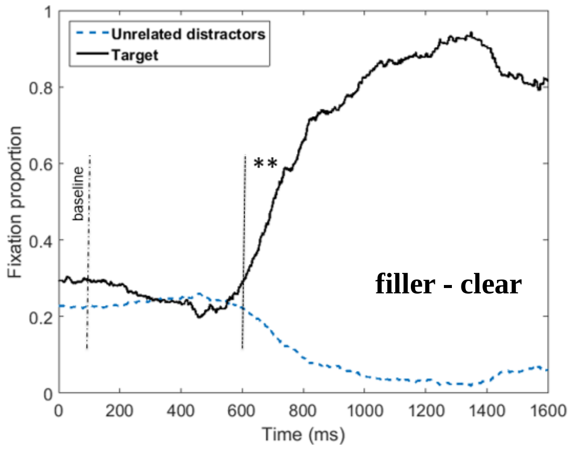

Figure 1: Fixation proportions for the target word and the average of the three distractors on filler trials in the clear. Plotted from the onset of the spoken target word until 1600 ms after spoken word onset.

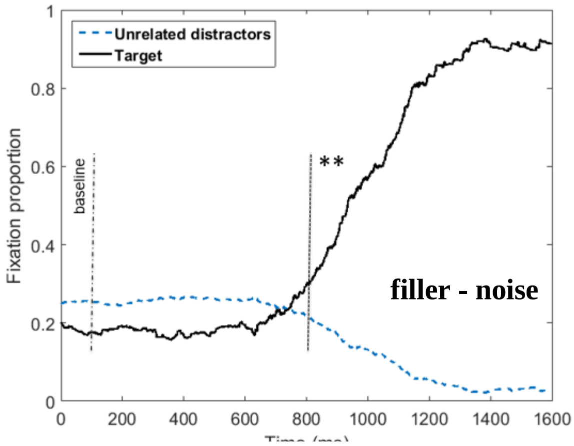

Figure 2: Fixation proportions for the target word and the average of the three distractors on filler trials in noise. Plotted from the onset of the spoken target word until 1600 ms after spoken word onset.

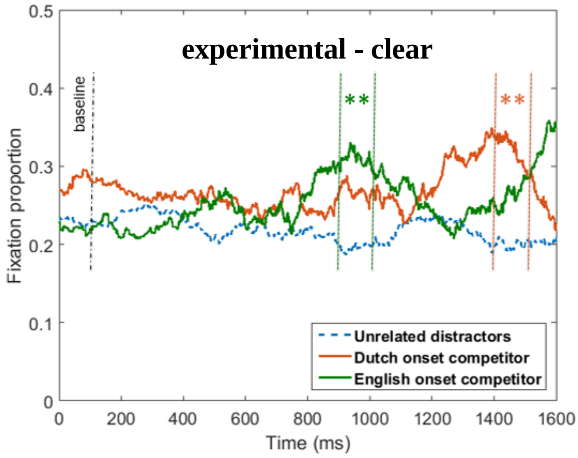

Figure 3: Fixation proportions for English competitor, Dutch competitor and the three average of the two distractors on experimental trials in the clear. Plotted from the onset of the spoken target word until 1600 ms after spoken word onset.

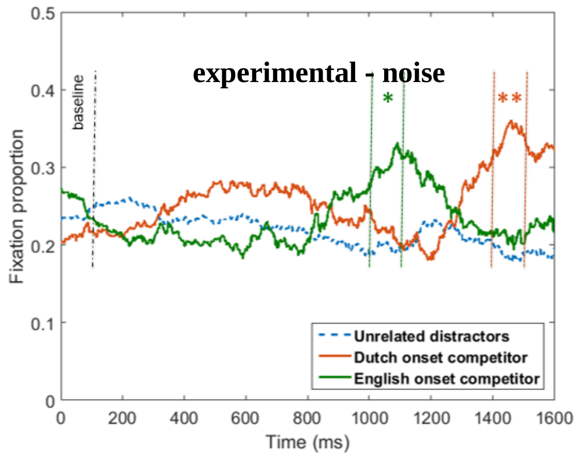

Figure 4: Fixation proportions for English competitor, Dutch competitor and the average of the two distractors on experimental trials in noise. Plotted from the onset of the spoken target word until 1600 ms after spoken word onset.

4. Discussion

Previous research has suggested that words, which overlap in phonological onset, compete for activation (Allopenna et al., 1998). Furthermore, it has been shown that when processing non-native speech, phonological competitors from both the native and non-native mental lexicon are activated, which results in increased competition for non-native listeners compared to native listeners (Weber & Cutler, 2004; Miriam & Spivey, 2003). Lastly, processing (native) speech in noise has been argued to result in increased phonological competition compared to processing speech in clear (Ben-David et al., 2011; Brouwer & Bradlow, 2016). The present study investigated the effects of both processing non-native speech and processing speech in noise, in order to determine the relative importance of the native and non-native mental lexicon. Participants were presented with an English spoken target word, as well as a display containing a Dutch (but not English) onset competitor, an English (but not Dutch) onset competitor and two unrelated distractors. The spoken target word was either presented in clear or masked by speech-shaped noise.

With regard to processing non-native speech, I found a bias for English phonological onset competitors (900 ms after word onset) and a bias for Dutch phonological onset competitors (1400 ms after word onset). With regard to processing non-native speech *in noise* I found a similar pattern. Firstly, English competitors were biased at around 1000 ms after word onset. This effect was only significant by subjects (not by items) and appeared approximately 100 ms later than in the clear condition. Secondly, Dutch competitors were biased at around 1400 ms after word onset. This bias appeared in the same time window as in the noise condition.

These results imply that, in the clear, there is competition between words from both the Dutch and English mental lexicons, confirming what previous studies have found (Weber & Cutler, 2004; Miriam & Spivey, 2003; Lagrou et al., 2013). Furthermore, these results imply that the relative importance of the English mental lexicon decreases once non-native English speech is processed in noise; the bias for English competitors was less strong and appeared later than in the clear condition. The observed effect of noise on the relative importance of the English mental lexicon when processing English speech supports the hypothesis that Dutch natives fall back onto their native language when processing English speech in noise. A possible interpretation of this finding could be that we tend to fall back onto processes we have more experience with when a certain task increases in difficulty. Some indication for such a mechanism comes from a study by Keysar, Hayakawa and An (2012) who showed that using a foreign language reduces decision-making biases in, for example, a gambling task. Their participants were less risk-seeking when presented with a bet in a non-native tongue, possibly because they were less emotionally attached and felt more ‘comfortable’ in their native tongue. In a similar vein, I speculate that in the present study the presence of background noise may have encouraged the use of native processing strategies. This is because native processing strategies are arguably far more automatized than their non-native counterparts due to greater experience.

The present study deviates from previous studies in that most previous studies observed phonological onset effects between 200 and 600 ms after spoken word onset. In the present study, I found effects between 900 and 1500 ms after spoken word onset. A possible explanation for this could be a combination of methodological choices: In contrast to previous studies, on experimental items in the present study the spoken target word was not presented as a picture on the screen. Moreover, participants were instructed to carry out an active task, that is to decide whether the spoken target word was depicted on the screen or not. It seems possible that participants waited until the offset of the spoken target before initiating their visual search. Future research needs to replicate these results. It would also be interesting to perform the same experiment, yet this time without an active task. A plain look-and-listen experiment could be performed to find out whether the nature of the task indeed had something to do with the delay of fixations to the competitors, as compared to previous studies.

5. Conclusion

When compared to the baseline (0-99 ms), there were significantly more fixations to the English competitor between 900 and 999 ms after word onset, and more fixations to the Dutch competitor between 1400 and 1499 ms after word onset when processing English speech in the clear. This is in line with the hypothesis that Dutch natives who process English speech experience competition from words of both their English and the Dutch lexicon. When background noise was added to the English speech, the fixations to the Dutch competitor remained more or less the same (also at 1400-1499 ms). The fixation to the English competitor appeared slightly later, at around 1000 to 1099 ms, and the effect was less strong in noise than in clear. Thus, the relative importance of the English lexicon decreases in noise, while at the same time Dutch natives put more emphasis on their native lexicon.

References

Allopenna, P., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, 38, 419-439.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., Schneider, B. A. (2011). Effects of Aging and Noise on Real-Time Spoken Word Recognition: Evidence From Eye Movements. *Journal of Speech Language and Hearing Research*, 54, 243-262.

Bobb, S. C., Drummond Nauck, L. Y., Altvater-Mackensen, N., Von Holzen, K., & Mani, N. (in press). Listening with your cohort: Do bilingual toddlers co-activate cohorts from both languages when hearing words in one language alone? In Schwieter, J. W. (Ed.) Cognitive control and consequences in the multilingual mind. Amsterdam: John Benjamins Publishing.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS ONE, 5*(5), e10773

Brouwer, S., & Bradlow, A.R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. doi:10.1007/s10936-015-9396-9

Dahan, D., & Gaskell, M.G. (2007). The temporal dynamics of ambiguity resolution: Evidence from spoken-word recognition. *Journal of Memory and Language*, 57, 4, 483-501.

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.

Huettig, F., & McQueen, J.M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57, 4, 460-482.

Keysar, B., Hayakawa, S.L., & An, S.G. (2012). The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases. *Psychological Science*, 23(6), 661-668.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013). Interlingual lexical competition in a spoken sentence context: Evidence from the visual world paradigm. *Psychon Bull Rev*, 20, 963-972.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behaviour Research Methods*, 44, 325-343.

Marian, V., Bartolotti, J., Chabal, S., Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE 7(8): e43230*.doi:10.1371/journal.pone.0043230

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, 2, 97-115.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71-102.

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

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165-178.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Saslow, M. G. (1967). Effects of Components of Displacement-Step Stimuli Upon Latency for Saccadic Eye Movement. *J. Opt. Soc. Am.* 57, 1024-1029.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50, 1-25.

Appendix: overview of the stimulus words

Dutch translations of distractors are given in brackets.

Item number	*Target word*	*English competitor*	*Dutch competitor*	*Distractor 1*	*Distractor 2*
1	wallet	wardrobe	wokkel	mouth (mond)	ball (bal)
2	mayor	maple	meetlint	cello (cello)	smoke (rook)
3	parlour	party	patat	hippo (nijlpaard)	nail (spijker)
4	carrot	carriage	ketting	toilet (toilet)	mattress (matras)
5	carving	casket	kapstok	peanut (pinda)	roof (dak)
6	dog	dot	dop	stamp (postzegel)	train (trein)
7	bag	bat	berg	cable (kabel)	painter (schilder)
8	brim	brick	bril	sausage (worst)	ski (ski)
9	bowl	bone	boom	dress (jurk)	tent (tent)
10	cattle	cap	kers	monkey (aap)	snow (sneeuw)
11	wink	wing	wip	mango (mango)	thermos (thermosfles)
12	wizard	window	wimpel	jeans (spijkerbroek)	bike (fiets)
13	rain	rail	reep	nurse (verpleegster)	mouse (muis)
14	map	match	mes	turtle (schildpad)	bed (bed)
15	pillow	pigeon	pinpas	jacket (jas)	honey (honing)
16	story	stomach	stormram	llama (lama)	table (tafel)
17	bandage	backpack	bergtop	wine (wijn)	sailor (matroos)
18	dentist	desert	deksel	shirt (overhemd)	pencil (potlood)
19	danger	daisy	deken	fish (vis)	pen (pen)
20	lobster	locker	lorrie	flute (fluit)	orange (sinaasappel)
21	lock	log	lont	hammer (hamer)	apple (appel)
22	soccer	saucer	sokkel	trumpet (trompet)	laptop (laptop)