Master's Thesis
in Language and Communication
(Research)

Perception of Native and Non-Native Lombard Speech by Native Speakers

Elisabeth Süß

February 11th, 2019

Supervisors:
Prof. Dr. Mirjam Ernestus
Katherine Marcoux (M.Sc.)

CLS | Centre for Language Studies
Radboud University

Abstract

Perception of Native and Non-Native Lombard Speech by Native Speakers

Elisabeth Süß

Even though many non-native (L2) speakers produce speech in noisy environments (so-called “Lombard speech”) regularly or even daily, the perception of L2 Lombard speech has not been studied yet. We studied how native (L1) and non-native listeners perceive L2 Lombard speech. We compared words in focus position in Lombard speech (NF) with words in non-focus position produced in quiet (QNonF).

Fifty-eight native or non-native English speakers determined whether the same keyword sounded more native-like when produced in NF or in QNonF. The keyword was produced twice in one of two possible orders: either first in NF and then in QNonF or the other way around. The listeners heard stimuli produced by eight L1 speakers of American English and eight L2 English speakers (L1: Dutch), blocked by speaker. The 28 keyword pairs consisted of three categories: words with initial /θ/ (e.g., *throne*), Dutch-English cognates with a schwa in American English and a full vowel in Dutch (e.g., *banana*), and words with final voiced obstruents (e.g., *club*).

Linear mixed effects modeling showed that interactions between the speaker nationalities and keyword categories significantly influenced whether the listener chose NF or QNonF. The general trend was that American (L1) speakers were perceived to sound more native in QNonF than NF and Dutch speakers (L2 speakers of English) showed the opposite pattern. This difference between NF and QNonF was particularly noticeable for the schwa category. For one of the orders of the two sound files, the theta category showed the opposite pattern compared to the other two keyword categories. Furthermore, the order of the sound files influenced the listeners significantly such that the second sound file was preferred irrespective of the order of the two sound files. In conclusion, listeners perceive the accentedness of native and non-native Lombard speech differently depending on the keyword category.

Table of Content

List of Tables

List of Figures

1 Introduction

1.1 Lombard Speech

When speakers have a conversation in a background of noise, they increase their vocal effort and fundamental frequency (f_0). Etienne Lombard (1911) was the first to report this type of speech and since then it has been referred to as “Lombard speech”. Lombard speech is not only louder than normal speech, but the f_0 and f_1 are also increased, segments are lengthened, and the spectral center of gravity is shifted upwards (Pisoni, Bernacki, Nusbaum, & Yuchtman, 1985). This effect of noise on one’s speech has been referred to as the “Lombard effect” (Junqua, 1993) and “Lombard reflex” (van Summers et al., 1988). In this thesis, an experiment on the perception of native versus non-native Lombard speech will be presented.

Several studies have shown that, when presented in noise, Lombard speech is more intelligible than normal speech (Dreher & O’Neill, 1957; van Summers et al., 1988; Pittman & Wiley, 2001; Lu & Cooke, 2008). The Lombard effect is enhanced by an increasing noise level as well as by an increasing number of competing speakers. In other words, the speaker’s vocal intensity increases as the energetic masking (i.e., signal degradation due to listening environments (Mattys, Brooks, & Cooke, 2009)) increases (Lu & Cooke, 2008).

After removing intensity differences, the Lombard benefit remains, so the Lombard benefit cannot be solely attributed to a higher vocal intensity (Junqua, 1993; Lu & Cooke, 2008). Shifting the spectral energy towards higher frequency regions can improve intelligibility in noise effectively and thus contribute to the Lombard benefit (Lu & Cooke, 2009). While placing information in regions that are less affected by noise, Lu and Cooke found that this was not the case in their study which compares speech produced in low-pass and high-pass noise. They hypothesize that the shifting of spectral energy towards higher frequencies might be linked to how relevant different frequency regions are in speech perception or monitoring of one’s own speech (Lu & Cooke, 2009).

1.2 Lombard Speech and Non-Nativeness

Most of the existing literature on Lombard speech studies the perception and production of L1 Lombard speech. However, multilingual speakers are the norm rather than the exception worldwide (Romaine, 1996). In the EU, 64.7% of 25- to 64-year-olds – the working population – speak at least one foreign language (Eurostat, Foreign lan-

guage learning statistics). Many of these Europeans use a foreign language to communicate in their everyday lives, which includes conversing in noisy environments. Consequently, L2 Lombard speech production and perception are omnipresent phenomena in multilingual environments within Europe and worldwide.

Most of the studies on L2 Lombard speech focus on non-native perception of L1 Lombard speech (e.g., Cooke & Lecumberri, 2012; Junqua, 1993). For example, Cooke and Lecumberri (2012) tested L2 listeners on a perception task and compared their results with data from L1 listeners who had performed the same task in an earlier study (Lu & Cooke, 2008). Both normal speech and Lombard speech produced by native speakers were presented in noise and in quiet. In noise, both native and non-native listeners recognized Lombard speech more accurately, but the Lombard benefit was slightly larger for the native group. Both listener groups seemed to have profited from the slower speech rate and the larger vowel space of Lombard speech, which are features that are shared across languages. In contrast, non-native listeners probably benefited less from language-specific features, such as vowel length and voicing contrasts of obstruents (Cooke & Lecumberri, 2012). In quiet, the result pattern was reversed for the non-native listeners, so they recognized normal speech more accurately than Lombard speech. Unfortunately, Lu and Cooke (2008) did not report on the quiet condition of the native listeners.

While the previously discussed papers study non-native perception of Lombard speech, only Li (2003) studied non-native production of Lombard speech. Recordings of Cantonese and English speakers reading English sentences in quiet and in 70 dB of cafeteria noise were presented to native speakers of English. These stimuli were presented with noise as well as without noise. They obtained intelligibility scores, comprehensibility ratings, and judgments on the degree of foreign accent for both groups of speakers. The transcription data from the intelligibility test showed that there was a Lombard benefit in the noise-masked condition for both native and non-native speakers, but this was not reflected in the accentedness ratings of the non-native speech. The accentedness ratings were similar in all listening conditions.

L2 speech differs from L1 speech in three major aspects: First, speakers usually have a foreign accent when speaking an L2 (Davies, 2015). Second, speaking in an L2 results in facing a higher cognitive load (van Summers et al., 1988). Third, L2 speakers may have to produce sounds that do not exist in their L1 phoneme invento-

ry. These differences demonstrate that results from research on L1 Lombard speech cannot be transferred to L2 Lombard speech.

Studying accentedness is relevant due to two main factors: First, strong foreign accents can lead to lower intelligibility (Langdon, 1999). Second, foreign accents may influence how we perceive others (including their competencies, Langdon, 1999). Thus, we focus on accentedness in our experiment on L2 Lombard speech in which we aim to answer the following research question:

> *"How do native speakers perceive native versus non-native Lombard speech in terms of accentedness?"*

In addition to the main research question, we also aim to answer a sub-question:

> *"How are different keyword categories (produced by native and non-native speakers) perceived by the listeners?"*

1.2.1 Experiment 1a and 1b

Experiment 1a was conducted to find eight Dutch speakers with an intermediate English proficiency for the main experiment (Experiment 2). These speakers were selected from 23 Dutch speakers who had been recorded for a production experiment. Dutch speakers with a moderate foreign accent in English represent the population we are studying the best. However, Dutch speakers with a strong foreign accent would not represent the average Dutch learner of English adequately and Dutch speakers with a slight foreign accent would sound too similar to English native speaker. Two American speakers from the same production experiment were also presented to establish a norm for native speakers and to identify listeners who do not judge native speech as being native (Jesney, 2004). The Experiment 1a was conducted online on LimeSurvey (LimeSurvey GmbH, Hamburg).

The accentedness rating pilot was conducted again (Experiment 1b) to receive accentedness ratings on normalized sound files because the volume of the sound files had not been normalized in Experiment 1a. The advantage of normalizing is that when sound files have been normalized, differences between trials cannot be due to differences in volume. Moreover, this enabled us to compare the accentedness ratings from Experiment 1a and 1b and consequently study the effect of normalization of the sound files.

Since Experiment 1b was conducted after the Dutch speakers had been selected, the speaker selection was not influenced by results from Experiment 1b. Due to time constraints, Qualtrics (Qualtrics, Provo, UT) was used for Experiment 1b be-

cause uploading sound files for five experimental lists can be done much faster in Qualtrics than in LimeSurvey (LimeSurvey GmbH, Hamburg). LimeSurvey has the advantage that sound files are played automatically at the beginning of each trial. This ensures that the participants only listen to the sound files (and thus follow the instruction to only play the sound files once). Ideally, Experiment 1a and 1b would have been conducted on the same website with the same procedure (either with sound files that were played automatically at the beginning of each trial or manually).

1.2.2 Experiment 2

In order to answer the main research question, we conducted a forced choice experiment in which native speakers of American English listened to Lombard speech and normal speech produced by native and non-native (L1: Dutch) speakers of English. Because the student population that was tested was so linguistically diverse, we included non-native English speakers as well. We decided to record American speakers and not British ones because the Dutch speakers' English sounds more American than British. For example, many of the Dutch speakers who were recorded produced flaps, which is characteristic of American English.

The stimuli consisted of words that pose difficulties for many Dutch speakers learning English. Words from three different keyword categories were presented: words with initial theta, words with a schwa in pre-stress position, and words with final voiced obstruents. First, words with initial /θ/ (theta) are difficult for Dutch learners of English because Dutch does not have /θ/ in its phoneme inventory. Second, the schwa keywords were Dutch-American cognates with a full vowel in pre-stress position in Dutch, but a schwa in pre-stress position in American English. Furthermore, the spelling of the words suggests that the letters represent a full vowel and not a schwa. These two reasons make the schwa keywords difficult words for Dutch speakers learning English. Third, keywords ending in voiced obstruents pose difficulties for Dutch learners of English because Dutch has final devoicing.

In every trial, the listener heard a keyword twice – once produced in quiet in off-focus position in the sentence (QNonF) and once produced in noise in focus position in the sentence (NF). The keywords had been recorded in four conditions: quiet non-focus, quiet focus, noise non-focus, and noise focus. Therefore, the two conditions, NF and QNonF, differ not only in whether the keyword is produced in quiet or in noise but also whether it is produced in focus position or in off-focus position. NF and QNonF are the extreme conditions out of the four conditions that were recorded be-

cause NF requires the highest effort from the speaker and QNonF requires the lowest one. For instance, in the NF condition, producing Lombard speech requires more effort than producing normal speech and speakers pronounce words in focus position more clearly which also leads to a higher effort. NF and QNonF were selected for Experiment 2 (the main experiment) despite potentially confounding differences in focus position because the other two conditions (noise non-focus and quiet focus) would show a pattern that is between the patterns of the two extreme conditions.

Half of the trials were presented in order 1 (NF – QNonF) and the other half in order 2 (QNonF – NF). The listener was asked to indicate which version of the word (QNonF versus NF) sounded more native to them. This enabled us to compare the conditions directly to each other instead of comparing ratings of the two conditions.

In a lab rotation, we previously piloted the experiment with eight Dutch speakers and two American speakers. In the result section, findings from the lab rotation will be compared to results from Experiment 2.

1.2.3 Hypotheses

1. Based on results from the lab rotation, we hypothesize that the listeners will perceive the accentedness of normal and Lombard speech differently depending on the speaker's nationality. More specifically, we predict - based on the results from the lab rotation - that the American speakers will show a larger difference in accentedness between normal and Lombard speech than the Dutch speakers (normal L1 speech will be judged as sounding more native than L1 Lombard speech).
2. Moreover, we hypothesize that the keyword categories might show different patterns in their perception because L2 speakers are aware of some difficult phonemes and phonological rules of the L2 which affect the production of the keywords. However, they are not aware of all these rules (Vokic, 2010).

2 Experiment 1a

2.1 Methods

2.1.1 Listeners

Six native speakers of American English (five females; age range: 24-28, mean age: 25.5) participated in Experiment 1a. All participants had been raised monolingually and knew neither Dutch nor German. Katherine Marcoux's and my friends and acquaintances were recruited as listeners for the experiment.

2.1.2 Stimuli

The stimuli were recorded as part of a production experiment which was conducted at the Center for Language Studies in Nijmegen, the Netherlands. Question-answer pairs were produced in quiet and in noise (e.g., "Did the child ask if the apple was sweet? No, she asked if the tomato was sweet", see appendix for complete list of question-answer pairs). Six sentences from each of the 25 speakers were presented. All sentences were the answer of the question-answer pair (e.g., "No, she asked if the tomato was sweet.").

Only sentences from the "quiet" conditions (QNonF and QF) were presented. The sound files were not normalized. For each speaker, the six sentences were chosen randomly from 72 possible sentences (2 quiet conditions x 36 keywords). Two sentences from each keyword category were selected and it was ensured that half of the sentences were from QNonF and the other half from QF. The sentences were blocked by speaker so that the listener could get used to the speaker. Within each block, the sentences were randomized.

2.1.3 Speakers

23 Dutch learners of English (only females, age range: 18 – 29 years, average age: 21.04 years) were presented. The other seven Dutch speakers from the production experiment had not given us consent to use their recordings in online studies. Recordings from two native speakers of American English (only females, age: 23 and 28 years, average age: 25.5 years) were included as control speakers. The Dutch speakers had taken the English LexTale (Lemhöfer & Broersma, 2012), which measures vocabulary knowledge in English using a lexical decision task. Their average score on this test was 66.70 (SD: 15.32).

2.1.4 Procedure

The listeners were tested on LimeSurvey (LimeSurvey GmbH, Hamburg). The participants were randomly assigned to one of the five lists which contained the same speakers in different orders. Within each block, the order of the sentences was also randomized. The sound files were played automatically at the beginning of each trial and were only presented once. By clicking on the button to go to the next trial, the sound file of the following trial was played automatically, so it was a self-paced experiment. The participants indicated how native the sentences sounded on a scale from 1 (“native-like”) to 7 (“very strong foreign accent”). They were instructed to rate each sentence individually and not the speaker in general. The ratings for the sentences were then averaged to obtain an accentedness rating for each speaker.

2.2 Results

The average accentedness ratings for the two American speakers were 1.00 and 1.03 (note that 1 is the minimum of the scale and represents “native-like”). The average of these two ratings was 1.02, which indicates that they were clearly judged as sounding native-like. All listeners rated the sentences produced by American speakers with a 1 or 2, which demonstrates that they can identify native speech and are reliable raters.

The average accentedness ratings for the Dutch speakers ranged from 2.44 to 5.94 (a higher value indicates a stronger foreign accent) (Figure 1). The average of all accentedness ratings for the Dutch speakers was 4.65 (SD=0.85), so 0.65 higher than the scale’s midpoint. Eight speakers whose averages were exactly on the median (4.69) or around it (range: 4.44 – 5.03) were chosen as speakers for Experiment 2. This selection was independent of the results from Experiment 1b because Experiment 1b was conducted after Experiment 2.

In order to compare listeners to each other, the ratings from each listener were also averaged. When averaged across all trials each listener rated (both Dutch and American speakers), listeners had averages between 3.76 and 5.03 (SD: 0.49). One listener used the scale from 1 to 6 and the other five listeners used the whole scale (from 1 to 7).

3 Experiment 1b

3.1 Methods

3.1.1 Listeners

Nine native speakers of American English (two females, six males, one “other”; age range: 19-32, mean age: 26.22) participated in the experiment. All participants had been raised monolingually and did not know Dutch. Katherine Marcoux’s and my friends and acquaintances were recruited as listeners for the experiment.

3.1.2 Stimuli

Except for the normalization to 70dB, the stimuli were the same as in Experiment 1a.

3.1.3 Speakers

The same speakers as in Experiment 1a were presented.

3.1.4 Procedure

The listeners were tested on Qualtrics (Qualtrics, Provo, UT). The procedure was the same as in Experiment 1a except that participants clicked on a button to hear the sound file and were instructed to only listen to it once. This difference in procedure was due to technical differences between the two online survey websites.

3.2 Results

The average accentedness ratings for the American speakers were 1.13 and 1.07 (note that 1 is the minimum of the scale and represents “native-like”). The average of these two ratings was 1.10, which indicates that the American speakers were clearly judged as sounding native-like. All listeners rated the sentences produced by American speakers with a 1 or 2, which shows that they can identify native speech and are thus reliable raters.

The average accentedness ratings of the Dutch speakers were between 2.30 and 4.65 (a higher value indicates a stronger foreign accent) (Figure 1). The average of all of these averages was 3.73 (SD=0.66), so relatively close to the scale’s midpoint, namely 4.

In order to compare Experiment 1a to Experiment 1b and to examine if the same speakers would have been chosen in both Experiment 1a and 1b, eight speakers from Experiment 1b with an intermediate foreign accent in English were chosen.

These eight speakers had average accentedness ratings which were on the median (3.96) or around it (range: 3.59 – 4.11).

In order to compare listeners to each other, the ratings from each listener were also averaged. When averaged across all trials each listener rated, listeners gave average accentedness ratings between 2.24 and 4.47 (SD: 0.64). One listener used the scale between 1 and 4, one used it between 1 and 5, and the remaining seven listeners used the whole scale.

3.3 Comparison of Experiment 1a and 1b

3.3.1 Comparison of the Dutch speakers

Speaker 18 was judged as having the weakest foreign accent in both Experiment 1a and 1b. The averages of the accentedness ratings for this speaker were similar across the two pilots: 2.44 in Experiment 1a, 2.30 in Experiment 1b. All other speakers received much higher ratings (indicating a stronger foreign accent) in Experiment 1a than in Experiment 1b. This is also reflected in the range of averages of all speakers (Experiment 1a: 5.94 - 2.44 = 3.50; Experiment 1b: 4.65 - 2.30 = 2.20) and of the eight speakers with intermediate foreign accents (Experiment 1a: 5.03 – 4.44 = 0.59; Experiment 1b: 4.11 – 3.59 = 0.52). The averages across all speakers also show this pattern (Experiment 1a: 4.65, SD=0.85; Experiment 1b: 3.73, SD=0.66).

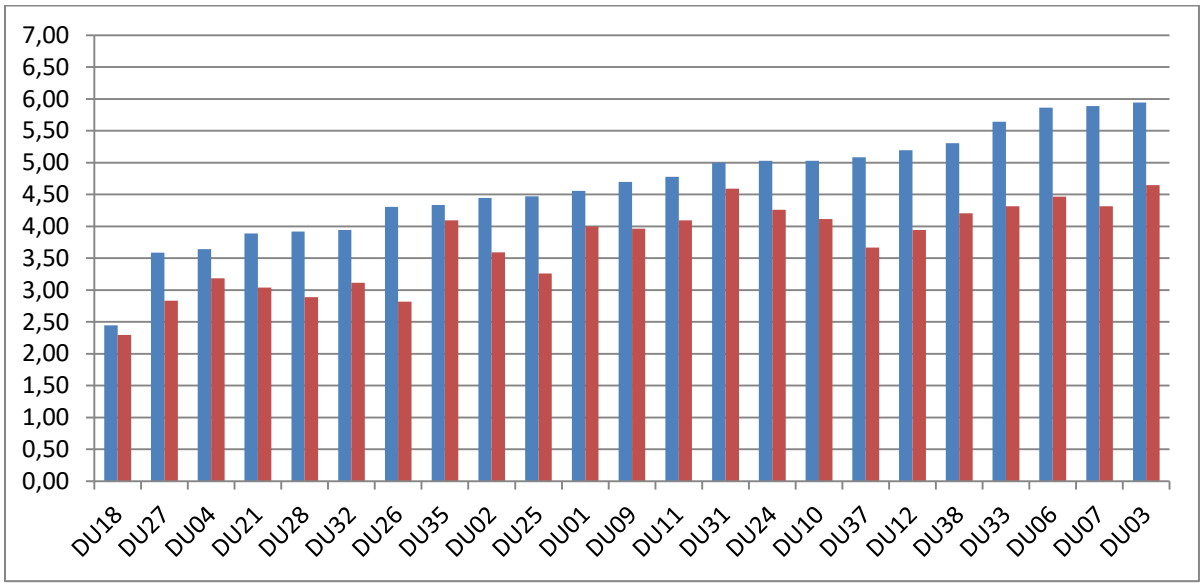

Figure 1: Experiment 1a and 1b: Average accentedness ratings for each speaker: Data from Experiment 1a in blue, data from Experiment 1b in red. The speaker number is on the x-axis. The average accentedness rating of the speakers is on the y-axis.

In the following, the two experiments will be compared in terms of which eight speakers had intermediate accents (and were thus selected for Experiment 2). Out of the 23 non-native speakers, the same seven speakers were judged as having the weakest foreign accents in both pilots (see Figure 1). Except for speaker 31, the same five speakers were judged as having the strongest foreign accent in both pilots. The remaining eleven speakers were among the eight speakers with an intermediate foreign accent in one of the pilots or in both. Five speakers were among the eight speakers with an intermediate foreign accent in both studies. The averages of speakers that were among these eight speakers in either Experiment 1a or 1b were in proximity to the averages of the eight speakers in the other pilot. The only exception for this pattern was speaker 31 who was among the eight speakers with an intermediate foreign accent in Experiment 1a and was the speaker with the second strongest foreign accent in Experiment 1b (see appendix for averages from both pilots).

3.3.2 Comparison of the Listeners

When averaging across all trials, listeners from Experiment 1a had an average accentedness rating between 3.76 and 5.03. Listeners from Experiment 1b had an average between 2.24 and 4.47, so the ratings from participants in Experiment 1b were generally less strict than ratings from Experiment 1a. This is also reflected in the part of the scale that listeners used: In Experiment 1b, a smaller percentage of the listeners used the whole scale (up to 7) for their ratings compared to Experiment 1a.

3.3.3 Effect of Normalizing the Volume

The normalization of the volume of the sound files in Experiment 1b might have influenced the accentedness ratings slightly. Some speakers received higher or lower ratings than in Experiment 1a. Changing the volume of sound files could have influenced how easily listeners notice a foreign accent. We hypothesize that when sound files are played at a louder volume level, listeners may judge sound files with a weak foreign accent more positively (more native-like) and sound files with a strong foreign accent more negatively (less native-like). Due to the very small sample size in both pilots, different results could also have been caused by differences between participants in Experiment 1a and Experiment 1b. However, all speakers expect speaker 31 were on the same part of the scale in both studies (weak foreign accent, intermediate

accent, strong accent), so the effect of normalizing the volume of the sound files did not change the overall result of Experiment 1a.

In conclusion, the two experiments have relatively similar results, but the listeners from Experiment 1a rated speakers as having stronger accents than listeners from Experiment 1b did.

4 Experiment 2

4.1 Methods

4.1.1 Listeners

Sixty students from the University of Alberta, Edmonton, Canada, participated in Experiment 2. One participant was excluded because they did not complete the second page of the questionnaire and another one was excluded because they were talking on their cell phone at the end of the experiment. The remaining 58 participants (34 females; age range: 18-27, mean age: 20.03) formed four groups. The first group consisted of 13 non-native speakers of English ("non-native" group). The other three groups consisted of native speakers of English but differed in whether the participants had been raised multilingually and whether they were familiar with languages that have final devoicing and/or do not have schwa ("problematic languages"). This familiarity was due to having learned the language, having lived in a country where this language is spoken for more than three months or speaking to non-native speakers of English who learned this language as their L1. The "monolingual and multilingual plus" group consisted of 18 listeners who had been raised monolingually or multilingually and who were not familiar with any of the problematic languages. The "monolingual minus" group consisted of 12 listeners who had been raised monolingually and were familiar with problematic languages. Finally, the "multilingual minus" group consisted of 15 speakers who had been raised multilingually and were familiar with problematic languages.

4.1.2 Stimuli

4.1.2.1 Keyword Categories and Keywords

The three keyword categories were words starting with /θ/ (e.g., *throne* /**θ**ɹoʊn/), words with final voiced obstruents (e.g., *club* /klʌ**b**/), and Dutch-English cognates with a schwa in pre-stress position (e.g., *balloon* /b**ə**ˈluːn/) in American English.

The category of keywords starting with /θ/ (theta) (voiceless "th") such as *throne* (/**θ**ɹoʊn/) is difficult for Dutch speakers of English because /θ/ is not part of the Dutch phoneme inventory. Consequently, many Dutch speakers substitute /θ/ with a /t/ /f/, or /s/ (Hanulíková & Weber, 2012).

The category of keywords ending in voiced obstruents (e.g., /d/) often poses difficulties for Dutch speakers of English because Dutch has final devoicing (Simon, 2010). This means that obstruents that are voiced when followed by a vowel within

the same word (e.g., *kinderen* /ˈkɪndərən/ “children”) are devoiced when they are in the final position in the word (e.g., *kind* /kɪnt/ “child”). Consequently, Dutch speakers often produce a /t/ instead of a /d/ or a /p/ instead of a /b/ at the end of an English word.

In the schwa category, the letter that corresponds to the schwa is often pronounced as a full vowel by Dutch speakers of English. In the Dutch version of the cognates, the vowel in the pre-stress position is a full vowel and not a schwa (e.g., [a] in [baˈlɔn]). In American English, the first vowel of *balloon* (/bəˈluːn/) is a schwa, but the word is spelled with an “a” in that position which might be interpreted as representing the full vowel [a]. Thus, Dutch speakers may produce the full vowel instead of the schwa in pre-stress position in these cognates.

	Number of words per category	**Standard American pronunciation**	**Dutch-accented pronunciation**
words with initial /θ/	9	/θ/, e.g., *throne* /**θ**ɹoʊn/	/t/, /s/, e.g., /**t**ɹoʊn/
words with voiced final obstruents	8	/b/, /d/, e.g., *club* /klʌ**b**/	/p/, /t/, e.g., /klʌ**p**/
cognates with schwa in pre-stress position	11	/ə/, e.g., *balloon* /b**ə**ˈluːn/	full vowel, e.g., /ɑ/ in /b**ɑ**ˈluːn/

Table 1: Keyword categories, number of words per category, examples of keywords, and Dutch-accented pronunciation of these keywords

The stimuli were elicited in a production experiment in which twelve keywords from each category were produced in each condition (see 4.1.2.2). From these 36 keywords, five had to be excluded because the American pronunciation of the word and the Dutch-accented pronunciation of the word constitute a (near) minimal pair (*theme – team, pub – pup, lab – lap, food – foot*). *Thermodynamics* was excluded because it often led to dysfluencies. *Thermometer* was excluded because it was very often produced with incorrect word stress. *Massage* was excluded because it was often substituted with *message*. After excluding these eight keywords from the original 36 ones, 28 keywords were suitable for Experiment 2 (see Table 1 for details).

4.1.2.2 Conditions and Carrier Sentences

The stimuli for Experiment 2 were taken from the same production experiment as the stimuli for Experiment 1a and 1b. For Experiment 2, keywords were segmented from

the question-answer pairs, while for Experiment 1a and 1b, the whole answer was segmented.

All keywords were produced in four different conditions: in focus in noise (NF), in focus in quiet, off-focus in noise, and off-focus position in quiet (QNonF). NF and QNonF were used for Experiment 2. In the focus conditions, the participants of the production experiment read contrastive question-answer pairs like this (see appendix for a complete list):

> "Did the family go to the **festival** in Barcelona? No, they went to the **parade** in Barcelona."

Words that are in contrastive focus are in bold. Participants were instructed to stress these bold words. The keywords are underlined. An example of a keyword in off-focus is *pub* in this contrastive question-answer pair:

> "Did **Bob** go to the pub in town? No, **Mary** went to the pub in town."

The keyword is both produced in the question and in the answer. We chose the instance from the answer for Experiment 2 because the keyword from the focus condition is also produced in the answer. For example, the keyword "parade" in focus position is only produced in the answer, not in the question.

4.1.2.3 Recording

The stimuli were recorded as part of a production experiment, which was conducted at the Center for Language Studies in Nijmegen, the Netherlands. The participants were recorded individually in a sound-proof booth wearing Sennheiser HD 215 MKII DJ headphones. They were recorded while they read question-answer pairs that were presented on a computer screen one at a time. The distance between the Sennheiser ME 64 or 65 microphone and the participant's mouth was 15 cm. All stimuli were produced in noise and quiet. In the noise condition, participants heard speech-shaped noise at 82 dB SPL via their headphones, which was used to make the participant produce Lombard speech. In the quiet condition, nothing was played via the participants' headphones.

4.1.2.4 Speakers

Sixteen female speakers were presented in Experiment 2. They had been recorded as part of the production experiment mentioned earlier. Eight of these women were native speakers of American English (age range: 19-28, mean: 22.13, SD: 2.67) and the other eight were native speakers of Dutch (age range: 18-24, mean: 20.75, SD:

1.85). The Dutch speakers were the eight speakers with an intermediate Dutch accent that were chosen in Experiment 1a. These non-native speakers of English had an average score of 64.03 (SD: 10.48) on the English LexTale (Lemhöfer & Broersma, 2012). All speakers had been raised monolingually by native speakers of the respective language. None of the speakers had a speech or hearing impairment. Some of the American speakers and all Dutch speakers had learned foreign languages.

4.1.2.5 Segmentation and Concatenation

For every speaker, I segmented the keywords from the carrier sentences from the NF and QNonF conditions using the word alignments by the Montreal Forced Aligner (MFA) (McAuliffe, Socolof, Mihuc, Wagner, & Sonderegger, 2017). The MFA uses a pronunciation dictionary, acoustic models, and written orthographic transcriptions of the sentences for the alignment. Katherine Marcoux created these transcriptions using a forced aligner. Speech and written transcriptions are aligned by using a pronunciation dictionary which maps graphemes to phonemes (McAuliffe et al., 2017). Dutch-accented English was added to the American English Carnegie Mellon University (CMU) pronouncing dictionary (i.e., final devoicing of obstruents, full vowels instead of schwas in pre-stress position in cognates, and /t/ and /d/ instead of word-initial /θ/).

I used a Praat script (Boersma & Weenink, 2018) to segment these words at the zero crossings closest to the word boundary. I listened to the resulting sound files one by one. When phonemes of the keyword were cut off or additional phonemes were audible in the sound file, I changed the boundaries by hand and moved them to the next zero crossings using the Praat function. Slips of the tongue that were still in the list of stimuli were removed manually. The sound files were normalized to 70dB and concatenated in both orders (order 1: NF – QNonF; order 2: QNonF – NF) with one second of silence in between using a Praat script (Boersma & Weenink, 2018).

4.1.3 Lists

The lists of Experiment 2 included the same 16 speakers with different ordering. The trials were blocked by speaker and the order of the trials within each block was created randomly, but constant across all lists. There were twelve different orders in which these blocks were presented which will be referred to as the “basic lists”. In each list, there were maximally three speakers from the same nationality in succes-

sion and maximally three keywords of the same category in succession in each block.

Since there were two possible orders for the two conditions to be presented in (order 1: NF – QNonF; order 2: QNonF – NF), half of the trials in each block had order 1 and the other half had order 2. To present all sound files in both orders in Experiment 2, every basic list was mirrored. This means that all trials with order 1 were replaced with order 2 and vice versa. Thus, a total of 24 lists were created (twelve basic lists x two orders).

4.1.4 Procedure

The participants were tested in a computer room at the University of Alberta in Edmonton. MB-QUART MBK C 800 headphones were used and Experiment 2 was completed online on WebExp (Webexperimenten van de Radboud Universiteit) on Think Center Lenovo computers using Windows 7.

The participants listened to the concatenated sound files which were presented without additional noise. They were instructed to press “Z” on the keyboard if the first version of the word sounded more native and “M” if the second version sounded more native. After four practice trials, the main experiment began, which contained 447 trials. Three breaks were distributed equally across the 447 trials. Lastly, the participants answered several demographic questions and a language background questionnaire (see appendix). Each experimental session lasted approximately 50 minutes.

4.1.5 Analysis

We performed a linear mixed effects analysis in `R` (version 1.1.463) (R Development Core Team, 2016) using `lme4` (Bates, Maechler, Bolker, & Walker, 2015) and `languageR` (Baayen, 2013). We studied how the keyword category, the speaker’s nationality, and the listener group affected the dependent variable “chosen condition” (the condition that sounded more native, QNonF versus NF). The order of the sound files was included as a fixed effect because there was a preference for the second sound file compared to the first one. This preference might show that either the right hand is preferred (possibly due to right-handedness). Alternatively, the second version of the word was more present in the memory of the listener because it was played last and is therefore preferred by the listener. Thus, “keyword category”, “speaker nationality”, “listener group”, and “order” were entered as the fixed effects of

the model. Listener, keyword, and speaker were the random effects of the model. The model was split up into an Order 1 Model and an Order 2 Model.

The model with the four-way interaction between “speaker nationality”, “keyword category”, “listener group” and “order” failed to converge. The Order 2 Model with the three-way interaction between “speaker nationality”, “keyword category”, and “order” failed to converge. Thus, the model with the three-way interaction was not selected for the complete data set. Because the three-way interaction did not converge, two-way interactions were included in the model. The fixed effects and two-way interactions were added to the model one after another. Akaike information criteria (AIC) were used to compare the fit of various models.

In the process of evaluating the significance of the random slopes, random slopes for “speaker nationality” were added first, then for “keyword category”, “listener group”, and “order”. This order was based on the theoretical importance of the fixed effects which decreased from “speaker nationality” to “order”. There were four random slopes in the model: by-listener random slopes for speaker nationality and order and by-speaker random slopes for the effect of keyword category and the effect of the listener group. Random slopes were removed from the model when the correlation between the random effect and the fixed effect was 0.80 or higher. These high correlations indicate that a model has been overparameterized (Baayen, Davidson, & Bates, 2008), which means the number of parameters of the model is higher than the estimated number of the parameters of the data (Upton & Cook, 2014).

The data was re-leveled: all keyword categories (theta, final devoicing, and schwa) and listener groups (monolingual and multilingual plus, monolingual minus, multilingual minus, non-native) were entered as the intercept one after another. Releveling all three models allowed comparisons between all keyword categories and all listener groups. While the tables in the result section of the models only show differences between final devoicing and schwa as well as final devoicing and theta, the releveling shows all possible comparisons. For example, the re-leveling also showed the difference between schwa and theta which is not included in the tables in the result section.

4.2 Results

4.2.1 Complete Model

The formula of the complete model was the following:

Complete_model=glmer (chosen_condition ~ (keyword_category * speaker_nationality) + (order * keyword_category) + listener_group + (1 | keyword) + (1 + speaker_nationality + order| listener) + (1+ listener_group + keyword_category| speaker), data=data13, family=binomial)

The data was binomial because the participants chose one of the two conditions in every trial depending on which one sounded the most native to them. All random effects of the model were significant. In all analyses, $p<0.05$ was considered significant (see description of Table 2 for significance codes). The final devoicing category, the Dutch speakers, order 1, and the "monolingual and multilingual plus" listener group were on the intercept.

	Estimate	SE	z value	Pr (>|z|)	
(Intercept)	0.419021	0.237364	1.765	0.07751	.
Keyword category: schwa	-0.100427	0.132244	-0.759	0.44761	
Keyword category: theta	-0.018465	0.152129	-0.121	0.90339	
Speaker nationality: EN	0.015668	0.225306	0.070	0.94456	
Order: order 2	-0.383247	0.096441	-3.974	7.07e-05	***
Listener group: monolingual minus	0.006357	0.274592	0.023	0.98153	
Listener group: multilingual	-0.370196	0.264289	-1.401	0.16130	
Listener group: non-native	-0.809517	0.292867	-2.764	0.00571	**
Keyword category: schwa x speaker nationality: EN	0.428378	0.143715	2.981	0.00288	**
Keyword category: theta x speaker nationality: EN	-0.308705	0.176765	-1.746	0.08074	.
Keyword category: schwa x order: order 2	-0.176843	0.071791	-2.463	0.01377	*
Keyword category: theta x order: order 2	0.132268	0.074942	1.765	0.07757	.

Table 2: Output of the glmer based on the complete data set: Dependent variable is the chosen condition (NF versus QNonF). Fixed effects and interactions are displayed with their corresponding estimate, standard error, z value, and significance value. Significance codes: ***=0.001; **=0.01; *=0.05; .=0.1

Table 2 shows that there was a main effect of order 2, which means that the choice of the listener was significantly influenced by the order of the sound files. The pattern can be summarized as a preference for selecting the second sound file as sounding the most native.

There was neither a main effect of speaker nationality nor of the keyword categories schwa and theta (see Table 2). However, there were significant interactions between schwa and American speakers as well as schwa and order 2. Differences between listener groups will be discussed in section 4.2.3. We can conclude that the decision of the listener was not influenced by the nationality of the speaker in and of itself. Re-leveling is necessary to understand how listeners are influenced by, for example, interactions between keyword category and speaker nationality.

4.2.1.1 Re-Leveling of the Model

Re-leveling of the model allowed comparisons between all keyword categories and all listener groups. For example, the re-leveling also showed the difference between schwa and theta, which is not included in Table 2 because final devoicing is on the intercept.

During the re-leveling, the model with “monolingual and multilingual plus”, the keyword category schwa, order 1 and Dutch speakers did not converge. Neither did the model with “monolingual minus”, the keyword category theta, order 1 and Dutch speakers, nor the model with “monolingual minus”, the keyword category schwa, order 1 and Dutch speakers.

The re-leveling showed that there was no significant difference between the final devoicing category and the theta category for Dutch speakers within the models which successfully converged. This comparison was not possible for American speakers because the models with American speakers and schwa or theta on the intercept did not converge.

When the keyword category schwa was on the intercept, there were significant interactions between American speakers and final devoicing as well as between the American speakers and theta. When theta was on the intercept, there were significant interactions between American speakers and schwa. There were significant differences between the keyword categories for the American speakers. For American speakers, both schwa (β=0.33, z=2.46, p<0.05) and theta (β=-0.33, z=-2.15, p<0.05) were significantly different from final devoicing, but the main effect of schwa was modulated by an interaction between schwa and Dutch speakers and between schwa and order 2. The main effect of theta was not modulated by an interaction. Neither the model with American speakers and schwa on the intercept nor the model with American speakers and theta converged. Thus, we do not know if theta and schwa

keywords were significantly different from each other when produced by American speakers.

In conclusion, while there was no significant main effect of keyword category for Dutch speakers, there were significant differences between keyword categories for American speakers. The interactions between speaker nationality and keyword category show that the speaker nationality influenced the schwa category significantly differently from the other two keyword categories. This can be seen in figures 2, 3, and 4 where the difference between the patterns of the two nationalities is much larger for schwa words than for the other two keyword categories.

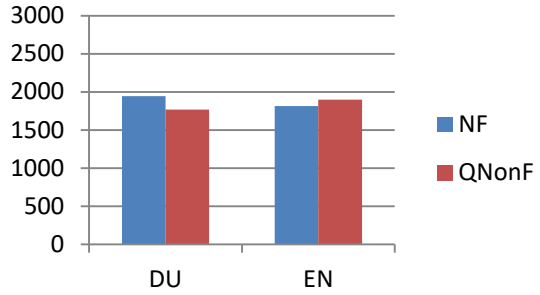

Figure 2: Data from trials with keywords with final voiced obstruents. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

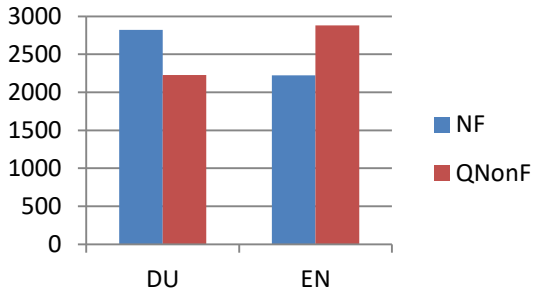

Figure 3: Data from trials with keywords with schwa in pre-stress position. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

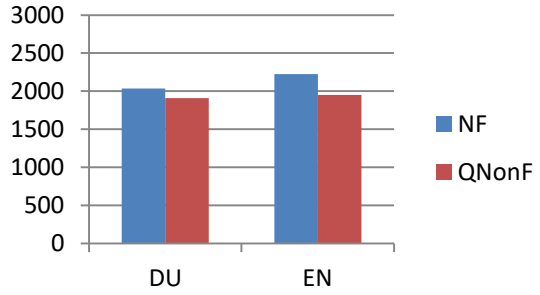

Figure 4: Data from trials with keywords with initial /θ/. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

When the keyword category schwa was on the intercept, there were significant interactions between order 2 and theta as well as order 2 and final devoicing. When theta was on the intercept, there was a significant interaction between order 2 and schwa. This shows that the order of the sound files influenced the schwa category significantly differently from the other two keyword categories. The difference between the

two conditions (NF and QNonF) was larger for the schwa category in both orders than it was for the other two keyword categories. The general preference for the second sound file as well as the larger difference between NF and QNonF for schwa keywords can be seen in Figure 5 and 6 again (QNonF preferred in order 1, NF preferred in order 2, see 4.2.1 for more information on the main effect of order).

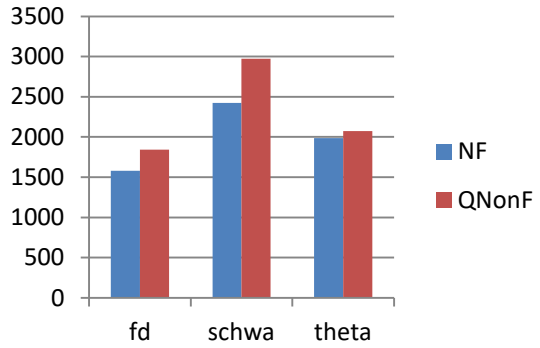

Figure 5: Data from trials with sound files in order 1. Keyword categories on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

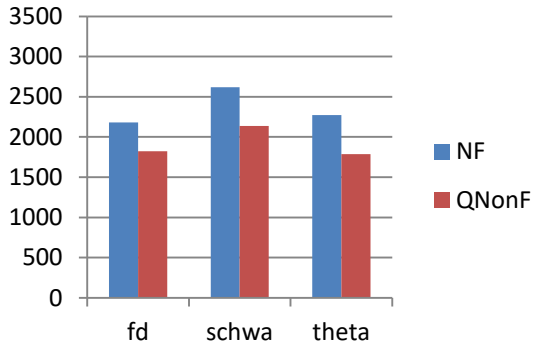

Figure 6: Data from trials with sound files in order 2. Keyword categories on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

4.2.2 Divided Data: Order 1 versus Order 2

Since there was a significant interaction between "order 2" and all three keyword categories the data was divided into two subsets, namely order 1 and order 2 (order 1: NF – QNonF; order 2: QNonF – NF). The complete model and the Order 1 and Order 2 Models differ in their random slopes because the latter models were based on an earlier version of the complete model.

4.2.2.1 Order 1 Model

The formula of the Order 1 Model was the following:

order1=glmer (chosen_condition ~ (keyword_category*speaker_nationality) + listener_group + (1|listener) + (1|keyword) + (1 |speaker), data=order1, family=binomial)

The final devoicing category, the Dutch speakers, and the "monolingual and multilingual plus" listener group were on the intercept. The Order 1 Model did not contain any random slopes for two reasons. First, the by-speaker random slope for the effect of order had been removed because the data had been split up into the two different orders. Second, the by-speaker random slope for the effect of keyword category had been removed because the correlation between by-speaker random slope and "keyword category" was very high (r=- .83).

	Estimate	SE	z value	Pr (>|z|)	
(Intercept)	0.41070	0.26993	1.522	0.128129	
Keyword category: schwa	-0.14504	0.14433	-1.005	0.314921	
Keyword category: theta	-0.15703	0.15128	-1.038	0.299275	
Speaker nationality: EN	0.07701	0.21766	0.354	0.723473	
Listener group: monolingual minus	0.32334	0.31747	1.018	0.308456	
Listener group: multilingual	-0.46370	0.29773	-1.557	0.119362	
Listener group: non-native	-1.07333	0.31131	-3.448	0.000565	***
Keyword category: schwa * speaker nationality: EN	0.45926	0.09988	4.598	4.26e-06	***
Keyword category: theta * speaker nationality: EN	-0.03440	0.10492	-0.328	0.742968	

Table 3: Output of the linear mixed effects model based on the “order 1” subset: Dependent variable is the chosen condition (NF versus QNonF). Fixed effects and interactions are displayed with their corresponding estimate, standard error, z value, and significance value. Significance codes: ***=0.001; **=0.01; *=0.05; .=0.1

Table 3 shows the fixed effects and interactions of the Order 1 Model. There was no main effect of speaker nationality, which means whether the speaker was Dutch or American did not influence the decisions (NF versus QNonF) of the listeners for the final devoicing keywords per se. For Dutch speakers, the final devoicing category did not differ significantly from the schwa category nor from the theta category. There was a significant interaction between American speakers and the keyword category schwa. Schwa keywords produced by American speakers showed a significantly different pattern from the other keyword categories produced by American speakers. The preference for QNonF compared to NF was significantly larger for schwa keywords produced by American speakers than for the other keyword categories produced by the same speakers (see Figure 8).

4.2.2.2 Re-Leveling of the Model

The model was re-leveled to be able to compare all keyword categories to each other and all listener groups to each other. For example, the re-leveling also compared the schwa and theta categories, which is not included in Table 3 because final devoicing is on the intercept. Additionally, it also showed whether the interaction between final devoicing and speaker nationality was significant.

The re-leveling of the data revealed the following: There was no significant difference between the schwa category and the theta category for the Dutch speakers (Figure 7). The two speaker nationalities only showed significantly different patterns from each other when schwa was on the intercept. Whenever schwa was on the intercept, the interactions between American speakers and final devoicing as well

as between American speakers and theta were significant. This interaction shows that the schwa category only showed a significantly different pattern than the other two keyword categories when the listeners had to judge speech produced by American speakers. When listening to keywords produced by American speakers, schwa words in QNonF were much more frequently judged as sounding more native than schwa keywords in NF (Figure 8).

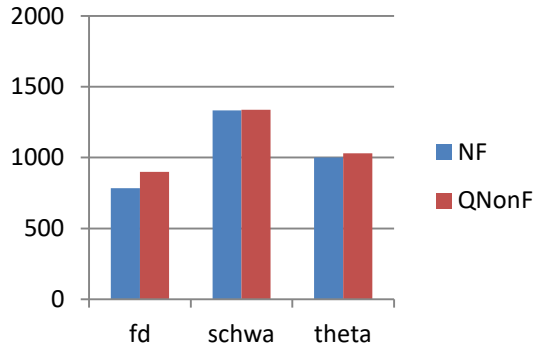

Figure 7: Order 1: Data from trials with Dutch speakers. Keyword categories on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

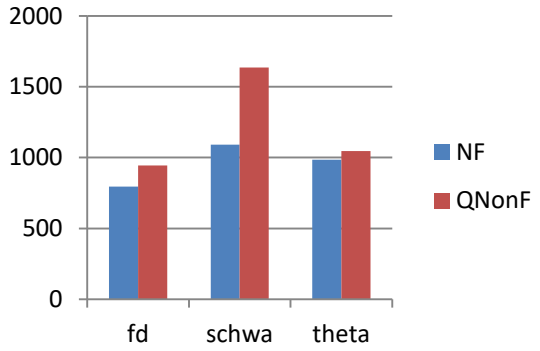

Figure 8: Order 1: Data from trials with American speakers. Keyword categories on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

4.2.2.3 Order 2 Model

The formula of the Order 2 Model was the following:

order2=glmer (chosen_condition ~ (keyword_category*speaker_nationality) + listener_group + (1|listener) + (1|keyword) + (1|speaker), data=order2, family=binomial

The final devoicing category, the Dutch speakers, and the "monolingual and multilingual plus" listener group were on the intercept. The Order 2 Model did not have any random slopes. First, the by-speaker random slope for the effect of keyword category was removed because it had also been removed from the Order 1 Model. Second, the by-speaker random slope for the effect of order was removed because the data had been divided into the two different orders.

	Estimate	SE	z value	Pr (>|z|)	
(Intercept)	-0.11589	0.25285	-0.458	0.64670	
Keyword category: schwa	-0.33723	0.10326	-3.266	0.00109	**
Keyword category: theta	0.18871	0.10784	1.750	0.08012	.
Speaker nationality: EN	0.23048	0.19593	1.176	0.23946	
Listener group: monolingual minus	0.36936	0.32291	1.144	0.25269	
Listener group: multilingual	-0.49837	0.30318	-1.644	0.10021	
Listener group: non-native	-0.79491	0.31690	-2.508	0.01213	*
Keyword category: schwa * speaker nationality: EN	0.49520	0.09825	5.040	4.66e-07	***
Keyword category: theta * speaker nationality: EN	-0.42899	0.10030	-4.277	1.89e-05	***

Table 4: Output of the linear mixed effects model based on the "order 2" subset: Dependent variable is the chosen condition (NF versus QNonF). Fixed effects and interactions are displayed with their corresponding estimate, standard error, z value, and significance value. Significance codes: ***=0.001; **=0.01; *=0.05; .=0.1

Table 4 shows that there was a main effect of schwa, but this main effect was modulated by the significant interaction between schwa and American speakers. Furthermore, there was a significant interaction between theta and American speakers.

4.2.2.4 Re-Leveling of the Model

The model was re-leveled to be able to compare all keyword categories to each other and all listener groups to each other. For example, the re-leveling also showed the difference between schwa and theta, which is not included in Table 4 because final devoicing is on the intercept.

The re-leveling showed that the keyword category schwa had a significantly different pattern from the other two categories. However, this main effect was modulated by the significant interaction between schwa and American speakers. Schwa keywords produced by American speakers were judged as sounding more native in QNonF, while the other two keyword categories produced by American speakers were perceived as sounding more native in NF. All three keyword categories were preferred in NF when they were produced by Dutch speakers, but this preference was particularly large for the schwa category.

The re-leveling showed that the interaction between American speakers and final devoicing was also significant. This suggests that all keyword categories showed a different pattern depending on the nationality of the speaker (see Figure 9 and Figure 10): The difference in perceived accentedness between NF and QNonF was small when native speech was presented, but NF was preferred when non-native speech had been presented. Schwa keywords produced by Dutch speakers were

preferred in NF, while they were slightly preferred in QNonF when produced by American speakers. The difference between NF and QNonF for schwa words produced by American speakers was significantly different from theta keywords (β=0.40, z=2.25, p<0.001), but not from final devoicing keywords. This main effect of schwa was modulated by the significant interaction between schwa and Dutch speakers.

Theta keywords were preferred in NF irrespective of the nationality of the speaker, but the preference for NF was much larger for American speakers. When American speakers were on the intercept, theta was perceived significantly different compared to the other two keyword categories (compared to final devoicing: β=-0.40, z=-3.93, p<0.001; compared to schwa: β=-0.24, z=-2.25, p<0.05), but significant interactions between theta and Dutch speakers modulated this effect.

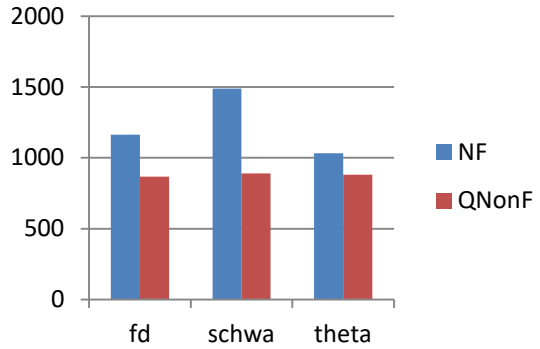

Figure 9: Order 2: Data from trials with Dutch speakers. Keyword categories on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

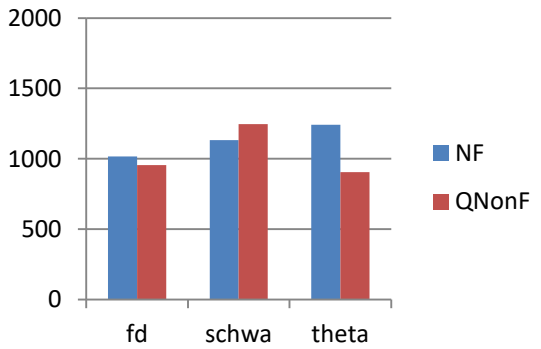

Figure 10: Order 2: Data from trials with American speakers. Keyword categories on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

Final devoicing and theta did not differ significantly from each other but approached significance (β=0.19, z=1.75, p=0.08), but this trend was probably modulated by the significant interaction between theta and American speakers (β=-0.43, z=-4.28, p<0.001). Figure 9 and Figure 10 show that only theta words produced by American speakers are judged as sounding more native in QNonF than in NF. All other keyword categories and speaker nationalities show the opposite pattern.

4.2.3 Listener Groups

The listener groups were re-leveled in all three models to allow comparisons between all groups, for example, “multilingual minus” versus non-native listeners. In all three models, non-native speakers showed significantly different patterns from the “monolingual and multilingual plus” and “monolingual minus” groups. The “multilingual minus” group showed different patterns in the three models.

4.2.3.1 Complete Model

In the complete dataset on which the complete model is based, the non-native speakers preferred NF compared to QNonF for both speaker nationalities. Their patterns for the two nationalities resemble one another strongly (Figure 11). In contrast to that, “monolingual minus” and “monolingual and multilingual plus” groups preferred QNonF compared to NF (Figure 13 and Figure 14). This preference for QNonF is stronger for American speakers than for Dutch speakers. These groups showed significantly different patterns: there was a main effect of non-native speakers when “monolingual and multilingual minus” (β=-0.81, z=-2.76, p<0.01) or “monolingual minus” (β=-0.82, z=-2.33, p<0.05) were on the intercept. The “multilingual minus” group did not differ significantly from any of the other groups. This group preferred NF compared to QNonF like the non-native listeners, but their preference for NF was smaller than of the non-native listeners (Figure 12).

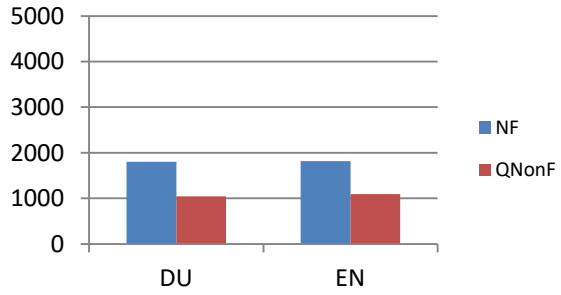

Figure 11: Data from the non-native listener group. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

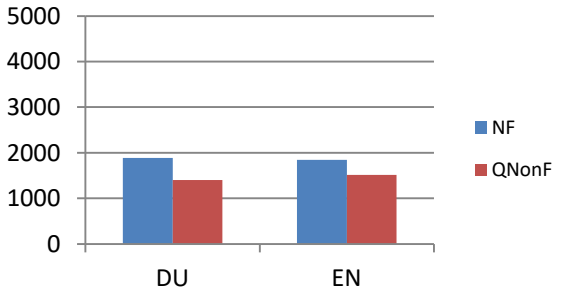

Figure 12: Data from the “multilingual minus” listener group. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on

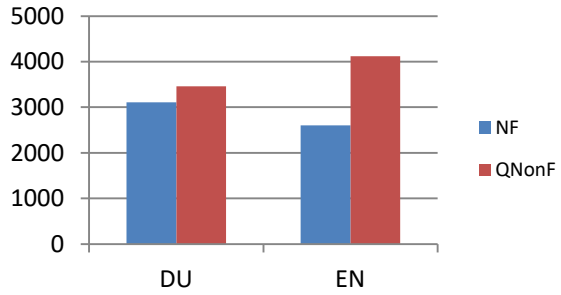

Figure 13: Data from the "monolingual and multilingual plus" listener group. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

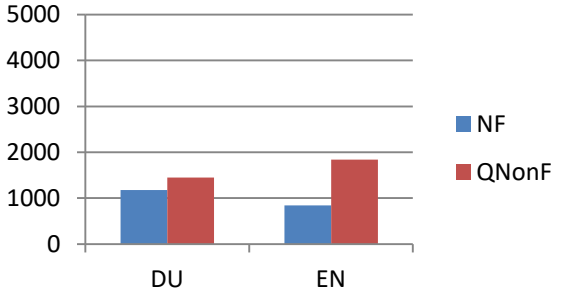

Figure 14: Data from the "monolingual minus" listener group. Dutch (DU) and American (EN) speakers on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

4.2.3.2 Order 1 Model

Similarly to the complete data set, when sound files were presented in order 1, non-native listeners judged NF as sounding more native than QNonF, whereas “monolingual and multilingual plus” and “monolingual minus” listeners preferred QNonF (Figure 15). These differences between the groups were significant: there was a main effect of non-native speakers when “monolingual and multilingual minus” (β=-1.07, z=-3.45, p<0.001) or “monolingual minus” (β=-1.40, z=-4.09, p<0.001) was on the intercept.

The non-native listeners only showed a significantly different pattern from “multilingual minus” listeners when Dutch speakers, “multilingual minus” listeners, and final devoicing or theta were on the intercept.

The “monolingual minus” group and the “multilingual minus” group only showed significantly different patterns from each other when “monolingual minus” was on the intercept or when Dutch speakers, schwa and “multilingual minus” listeners were on the intercept.

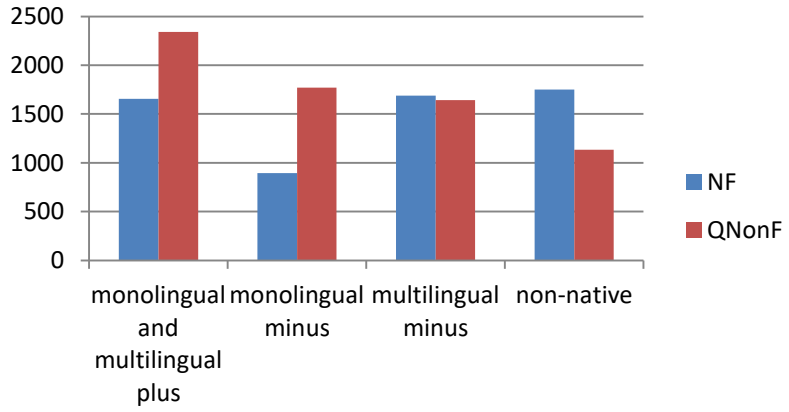

Figure 15: Order 1: listener groups. Listener groups on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

4.2.3.3 Order 2 Model

Just like in the complete data set and the order 1 subset, in the order 2 subset, the “multilingual and monolingual plus” group and the “monolingual minus” group showed a significantly different pattern from the non-native group (Figure 16). There was a significant main effect of non-native speakers when the “multilingual and monolingual plus” group (β=-0.79, z=-2.51, p<0.05) or the “monolingual minus” group was on the intercept (β=-1.16, z=-3.34, p<0.001). While the non-native listeners and the “monolingual minus” listeners showed the same preferences as in the previously discussed models, “multilingual and monolingual plus” preferred NF slightly more than QNonF.

Only the “monolingual minus” group showed a significantly different pattern from the “multilingual minus” group. These two groups showed the opposite prefer-

ence: “monolingual minus” preferred QNonF compared to NF, while “multilingual minus” preferred NF compared to QNonF.

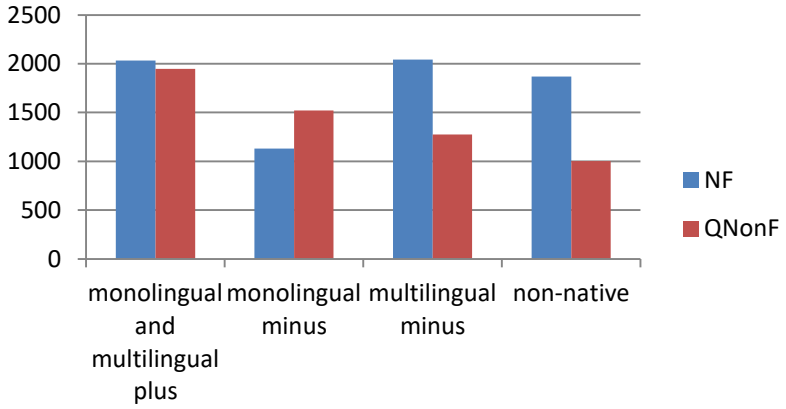

Figure 16: Order 2: Listener groups. Listener groups on the x-axis, total amount of times each condition (NF and QNonF) was judged as sounding more native than the other condition on the y-axis.

4.2.3.4 Listener Groups: Conclusion

Based on all three models, we can conclude that non-native speakers of American English judged normal and Lombard speech differently than participants who grew up monolingually speaking English (and possibly additional languages in case of “monolingual and multilingual plus”). Non-native speakers judged NF as sounding more native than QNonF irrespective of the order of the two sound files. The “monolingual minus” group judged QNonF as sounding more native than NF in both orders, but this preference for QNonF was larger in order 1 (where QNonF was the second sound file). The “monolingual and multilingual plus” group preferred QNonF compared to NF in order 1, but preferred NF slightly compared to QNonF in order 2. This shows how strongly the “monolingual and multilingual plus” group was affected by the order of the sound files. The “multilingual minus” group was also affected relatively strongly by the order of the sound files: in order 1, this group judged NF and QNonF as sounding equally native, while in order 2, NF was judged as sounding more native than QNonF. This pattern for order 2 is in line with the preference for the second sound file.

4.2.4 Comparison between Experiment 2 and the Lab Rotation

The results from Experiment 2 can be compared to the results from the Lab Rotation experiment because the Lab Rotation experiment was the pilot experiment for this thesis. Both in the lab rotation and the thesis (complete dataset and “monolingual and multilingual plus” listener subset), QNonF was judged as sounding more native compared to NF when sound files from American speakers were presented.

The pattern for Dutch speakers differed between the datasets: the subset from the thesis (“monolingual and multilingual plus”) and the lab rotation showed a very similar pattern, whereas the complete data set from this thesis showed a different pattern. The pattern from the subset and the lab rotation resemble each other (QNonF is slightly preferred compared to NF), probably because the requirements for the listeners were the same. However, in the complete data set from the thesis, NF was preferred compared to QNonF. Speaker nationality was a significant main effect in the lab rotation, whereas these preferences were only trends in the thesis.

The pattern of the three keyword categories also differed between the two studies. In the lab rotation, QNonF was always preferred compared to NF except for /θ/ produced by Dutch speakers. In Experiment 2 of this thesis (data from all listeners), NF was always preferred compared to QNonF when sound files from Dutch speakers were presented. When sound files from American speakers were presented, QNonF was preferred compared to NF in the final devoicing condition and in the schwa condition, but not in the theta condition. In the data from the “monolingual and multilingual plus” listener group, QNonF is preferred compared to NF for both speaker nationalities. The pattern of this subset resembles the lab rotation result. In both experiments, the theta condition showed a different pattern from the other two conditions, but in the lab rotation it was theta produced by Dutch speakers and in the thesis, it was theta produced by American speakers. In the lab rotation, the main effect of theta approached significance, while there were significant interactions with theta in the thesis.

Possible reasons for these differences are the following:

1. The lab rotation was based on data from ideal listeners (listeners who would be categorized as “monolingual and multilingual plus” in this thesis), whereas Experiment 2 of this thesis had various listener groups including non-native and “multilingual minus” listeners.
2. The experiment for the lab rotation was conducted online with volunteers, whereas Experiment 2 was conducted in a computer room, supervised by researchers.
3. There were fewer trials in the experiment for the lab rotation than in Experiment 2 of this thesis (83 to 85 trials depending on the experimental list versus

443 trials). This higher number of trials in Experiment 2 of the thesis led to larger statistical power.

4. Different Dutch speakers were selected for the two experiments and the selection methods differed. In the lab rotation, Dutch speakers were selected based on LexTale scores in English (Lemhöfer & Broersma, 2012) and performance on a perception task with casual American speech. Three different proficiency levels were selected, but the differences between them were not significant. In this thesis, the Dutch speakers with an intermediate foreign accent in English were selected based on accentedness ratings (see Experiment 1a).
5. The stimuli were cut again for Experiment 2 of the thesis which means that the sound files presented in the two experiments were not exactly the same. The final devoicing and the theta categories could have especially been affected by different segmentations because they are in the word-initial and word-final position.

"Order" as an additional fixed effect in the lab rotation model

When "order" was added as a fixed effect in the model from the lab rotation and a random slope was removed, the formula of the model was the following:

data_order.glmer=glmer (chosen_condition ~ Nativeness_of_speaker + Category_of_word + Order_of_soundfiles + (1|Word) + (1|ppt) + (1|Speaker_Number), data=data, family=binomial)

The by-speaker random slope of keyword category was removed because the correlation between "speaker" and "keyword category" was very high (theta: r=-0.87, schwa: r=-0.79).

There was a significant main effect of order (β=-0.46, z=-5.745, p<0.001). This shows that both in the lab rotation and in Experiment 2 of the thesis, the order of the two sound files affected the response of the listeners significantly.

Irrespective of whether "order" was part of the model, "theta" showed a significantly different pattern from "final devoicing" (model with "order": β=-0.50, z=-2.42, p<0.05) and there was a significant main effect of speaker nationality (model with "order": β=-0.94, z=-3.65, p<0.001). This is in stark contrast to Experiment 2 from this thesis, but might be due to differences in statistical power and the fact that there were significant interactions in Experiment 2 which involved keyword categories and the speaker nationalities.

4.3 Discussion

4.3.1 Interpretation of the Results

4.3.1.1 Main research question

In Experiment 2, we addressed the question of how native and non-native Lombard speech is perceived by native listeners. In the complete model with data from all listeners and final devoicing on the intercept, there was no significant main effect of speaker nationality. In general, that means that the participants did not judge the two conditions (NF and QNonF) differently depending on the nationality of the speaker in the final devoicing category. However, the significant interaction in the complete model between speaker nationality and the keyword category schwa showed that the speaker nationality influenced the decision of the listeners in the schwa category differently than in the other two categories. Listeners preferred NF for Dutch speakers and QNonF for American speakers. These results suggest that native Lombard speech sounded less native (with this speaker group “natural” seems more appropriate than “native”) than native normal speech. Native Lombard speech sounds probably less natural because of changes in formants, vowel durations, and so forth. However, non-native Lombard speech was judged as sounding more native than non-native normal speech. A possible reason for this might have been that the speakers’ effort was higher in FL than in QNonF, so the speakers articulated the keywords more carefully and therefore more native-like. These differences between NF and QNonF in the two speaker nationalities were larger for the schwa category than for the other categories. Differences between schwa and the other two keyword categories will be discussed in 4.3.1.2.

The Order 1 Model showed this pattern as well, namely that schwa was affected differently by speaker nationality than the other two categories. All keyword categories showed only slight differences between NF and QNonF when produced by Dutch speakers. However, when produced by American speakers, QNonF was preferred for all keyword categories, but this preference was bigger for schwa words than for the other keyword categories.

In the Order 2 Model, the perception of all three keyword categories was differently affected by speaker nationality. When Dutch speakers were presented, NF was preferred for all keyword categories (again this preference was larger for schwa words than for other keyword categories). When American speakers were presented, schwa words showed the opposite pattern and final devoicing words showed only a

small difference between NF and QNonF. Theta words produced by American speakers showed a much larger preference for NF compared to QNonF than theta words produced by Dutch speakers.

When comparing the Order 1 Model to the Order 2 Model, it became evident that sound files in order 1 produced by Dutch speakers and sound files in order 2 produced by American speakers showed relatively small differences between NF and QNonF. In contrast to that, sound files in order 2 produced by Dutch speakers and sound files in order 1 produced by American speakers showed large differences between NF and QNonF. However, not all interactions between speaker nationality and keyword category were significant. Whenever there were large differences between NF and QNonF, the second sound file was preferred, which is in line with the main effect of order in the complete model. It is surprising that the larger differences between NF and QNonF are observed once for Dutch speakers (order 2) and once for American speakers (order 1).

These interactions between speaker nationality and keyword category show that the speaker nationality influenced the decision of the listener, but that the keyword category modulated this effect. Hence, based on the trials from all listeners, we can confirm one aspect of Hypothesis 1, namely that the speaker groups show different patterns. Hypothesis 1 states that American speakers will show a larger difference in accentedness between normal and Lombard speech than the Dutch speakers will (normal L1 speech will be judged as sounding more native than L1 Lombard speech). We can confirm that listeners judged normal and Lombard speech produced by native and non-native speakers differently, but the keyword category modulated this effect. However, there was no significant main effect of speaker nationality. The difference that we had hypothesized between Dutch and American speakers was not observed in Experiment 2. Visual inspection of the data suggests that the differences between NF and QNonF were even smaller for American speakers than for Dutch speakers.

In contrast to the complete data set, the data from the “monolingual and multilingual plus” group, the ideal listener group for this experiment, showed exactly the pattern that was described in Hypothesis 1. This was probably due to the very similar selection criteria for the lab rotation and the “monolingual and multilingual plus” group because Hypothesis 1 was based on results from the lab rotation. QNonF was preferred compared to NF for both speaker nationalities, but especially for the American

speakers. We can conclude that native listeners prefer normal speech over Lombard speech for non-native speakers (Dutch speakers) in terms of accentedness and that this preference is even stronger for native speakers (American speakers).

The complete model showed that the order of the sound files influenced the choice between the conditions significantly. Listeners preferred the second sound file in both orders. Due to interactions between speaker nationality and order as well as keyword category and order, the model was divided into the Order 1 Model and the Order 2 Model to examine the differences between the orders in more detail.

4.3.1.2 Sub-question: Keyword categories

In the complete model and in the Order 1 Model, there was no main effect of keyword category. In the Order 2 Model, the keyword category schwa behaved significantly differently from the other two categories. However, this main effect was modulated by the significant interaction between schwa and American speakers. We can conclude that there were no significant differences between the keyword categories for the Dutch speakers, but significant interactions between keyword categories and speaker nationalities. Theta keywords and final devoicing keywords produced by American speakers were significantly different from each other. Schwa keywords produced by American speakers were also significantly different from final devoicing keywords, but this effect was modulated by two interactions.

In the Order 1 Model, schwa was affected differently by the speaker nationality than the other two keyword categories. In the Order 2 Model, all keyword categories were affected differently by the speaker nationality. Hence, we can only confirm Hypothesis 2 to some extent. Hypothesis 2 states that the two speaker groups will show a different pattern across the three keyword categories. In our results, differences between keyword categories were modulated by differences between speaker nationalities.

The differences between the schwa category and the other keyword categories in order 1 could have been caused by the following: First, this category was about differences in vowels and not in consonants. Second, this was the only word-internal category. These differences could have led to differences in production and/or perception of normal speech and Lombard speech. Differences in salience were possible between vowels and consonants and/or between word-internal phonemes and word-initial/word-final phonemes. Differences in the effect of noise on the production of vowels versus consonants could also have caused the differences. The

schwa category consisted of eleven keywords, while the other two categories consisted of eight and nine keywords. Consequently, there were more trials with schwa keywords (175 trials) than with theta (140 trials) or final devoicing keywords (128 trials) in the experiment. This means that the statistical power was higher for the schwa category than for the other two categories, so the chances of finding significant effects for schwa were higher than for the other two keyword categories.

The interaction between schwa and American speakers shows that listeners perceived the largest difference between normal and Lombard speech produced by Dutch and American speakers in the schwa category. This might also reflect that the production of schwa words was more affected by noise than the other two conditions. This stronger effect of noise (in perception, production, or both) led to opposite patterns of the two nationalities in the schwa condition.

When sound files were presented in order 2, not only schwa keywords showed a different pattern compared to the other two categories, but final devoicing and theta were also affected differently by speaker nationality. The theta category showed a surprising pattern, namely that theta words produced by American speakers were preferred in NF compared to QNonF. This preference was also observed for final devoicing keywords produced by American speakers, but in a much smaller magnitude than for theta keywords. This pattern within the theta category is surprising because QNonF is preferred for American speakers in most of the other keyword categories in both order 1 and order 2.

4.3.1.3 Listener groups

In all three models, non-native speakers showed significantly different patterns from the “monolingual and multilingual plus” and “monolingual minus” groups. Non-native speakers judged NF as sounding more native than QNonF irrespective of the order of the two sound files, while “monolingual minus” and the “monolingual and multilingual plus” group preferred QNonF compared to NF. This result for the non-native listeners is not in line with the results from Cooke and Lecumberri (2012) who found that non-native listeners perceive Lombard speech less accurately than normal speech. One obvious crucial difference between these two studies is that Cooke and Lecumberri studied intelligibility and we studied accentedness. We can conclude that non-native speakers of English judged normal and Lombard speech differently from participants who grew up monolingually speaking English (and possibly additional languages). The “multilingual minus” group showed different patterns in the three models.

4.3.2 Shortcomings of the Experiments and Future Research

The shortcomings of Experiment 1a and 1b are the small sample size and the fact that two different websites were used for the studies because of time constraints. Due to the difference in websites, the procedure between the two studies differed which makes it more difficult to compare their respective results (Experiment 1a: sound files were played automatically in each trial, Experiment 1b: participants had to click on a button to hear the sound file). In Experiment 1b, participants could have listened to the sound files multiple times even though they were instructed not to do so.

In Experiment 2, there should have been more listeners in the “monolingual and multilingual plus” group or a greater number of listeners in each group to have sufficient statistical power in the analysis. Ideally, all 58 listeners would have been members of the “monolingual and multilingual plus” group. Since there were relatively few listeners in this group, more listeners will be tested in the upcoming months to increase the statistical power.

Furthermore, it would be interesting to replicate the experiment using a scale instead of a binary forced choice to examine whether the results would be the same. The listeners would listen to one keyword in every trial and rate the foreign accent on a scale similar to the scale in Experiment 1a and 1b. We would explicitly instruct the participant to use the whole scale. The dependent variable would be the value on the scale and the condition (NF and QNonF) would be a fixed effect in the model. If listeners use the whole scale, the results could be more nuanced – at least for keywords for which the difference in accentedness between NF and QNonF is very large. I predict to find larger differences between NF and QNonF for some keywords or some speakers, but the overall pattern will probably be relatively similar. For trials with small differences between NF and QNonF, the binary forced choice is more suitable than a scale because a binary forced choice task prevents participants from rating both conditions with the same value.

In addition to accentedness of non-native Lombard speech, intelligibility of non-native Lombard speech should be studied more in future research. Both accentedness and intelligibility are important factors in non-native communication because they both influence whether the listener understands or misunderstands the speaker. Furthermore, accentedness and intelligibility influence how high the listening effort is.

In a future experiment, intelligibility of native and non-native Lombard speech and normal speech could be studied by presenting minimal pairs to native and possibly non-native listeners. The task of the listener would be to indicate if they have heard the standard American pronunciation (e.g., “theme”) or the Dutch-accented pronunciation (e.g., “team”) which forms a minimal pair with the standard American pronunciation. This design would be possible with the final devoicing category (e.g., “pub” versus “pup”) and the theta category (e.g., “theme” versus “team”), but not with the schwa category (e.g., “cadaver” with schwa versus a full vowel in the pre-stress position). Whether the listeners choose “theme” or “team” would show whether the intended word (“theme”) was perceived or the other member of the minimal pair (“team”). The frequency of all stimuli would be included in the analysis. If the intended word was chosen in the majority of the trials, this would demonstrate that the speaker is intelligible – at least with respect to theta in word-initial position and final devoicing. If the other member of the minimal pair was chosen in the majority of the trials, this would demonstrate the opposite. It would be interesting to compare speaker nationalities and keyword categories in this intelligibility study as well as to be able to compare the results to those from Experiment 2. I would expect to find a main effect or at least an interaction with speaker nationality in this intelligibility study, so I predict to find a difference between conditions (NF and QNonF) depending on the speaker nationality. This would be in line with the results from the lab rotation and Experiment 2 which also find a main effect of speaker nationality or an interaction between speaker nationality and keyword category.

5 Summary

In this thesis, we addressed the question of how native speakers perceive native versus non-native Lombard speech in terms of accentedness. Two accentedness rating studies (Experiment 1a and 1b) and a forced choice perception experiment (Experiment 2) were conducted to answer this research question.

In Experiment 1a and 1b, native speakers of American English rated the strength of the foreign accent in individual sentences produced by Dutch learners of English as well as two American speakers as a control. The participants used a 7-point scale to indicate the strength of the foreign accent and the average of these scores was computed for each speaker. Dutch speakers with a score on the median or close to it were selected as Dutch speakers for Experiment 2.

In Experiment 2, native and non-native speakers of American English listened to keywords produced in focus position in noise (NF) and in non-focus position in quiet (QNonF). Non-native listeners were included because the participants were university students in Canada and this group contained much linguistic diversity. The keywords were produced by eight native speakers of American English and eight non-native speakers of English (L1: Dutch). In every trial, the listener indicated whether the same keyword sounded more native in NF or in QNonF, which allowed a direct comparison of the two conditions. These keyword pairs consisted of three categories: words with initial /θ/ (e.g., *throne*), Dutch-English cognates with a schwa in American English and a full vowel in Dutch (e.g., *banana*), and words with final voiced obstruents (e.g., *club*).

In the complete data set of Experiment 2, there was no significant main effect of speaker nationality, but there was a non-significant trend that listeners preferred NF compared to QNonF for Dutch speakers and QNonF compared to NF for American speakers.

However, the result is different when we look at the “monolingual and multilingual plus” group. The participants in this group are the optimal listeners for Experiment 2 because they are not familiar with languages that have final devoicing and/or not schwa. The data from this listener group showed that native listeners prefer normal speech over Lombard speech for non-native speakers (Dutch speakers) and that this preference was even stronger for native speakers (American speakers).

The order in which the two conditions were presented influenced the decision of the listeners significantly, namely whether the same keywords sounded more na-

tive in NF or in QNonF. Sound files in the second position were judged as sounding native more often than the sound files in the first position. Since there was a significant interaction between the order of the sound files and the keyword category, the data set was divided into two data sets (order 1: NF – QNonF, order 2: QNonF– NF).

When sound files were presented in order 1, schwa showed a larger difference between NF and QNonF for the two speaker nationalities than the other two keyword categories did. In order 1, schwa keywords produced by Dutch speakers were judged as similarly native in both NF and QNonF, whereas schwa words produced by American speakers were judged as sounding much more native in QNonF.

When sound files were presented in order 2, all three keyword categories were affected differently by the speaker nationality. In order 2, final devoicing keywords and schwa keywords were judged as sounding much more native in NF than in QNonF. However, this preference was much smaller for final devoicing keywords produced by American speakers and NF was even preferred slightly more than QNonF for schwa keywords. In contrast to these two keyword categories, theta keywords produced by American speakers were judged as sounding much more native in NF than in QNonF. Thus, the theta keywords produced by American speakers in order 2 show a pattern like the keywords produced by Dutch speakers in order 2.

Non-native listeners were also included because the student population that was tested was linguistically diverse. In all three models, the “monolingual and multilingual plus” and “monolingual minus” groups showed significantly different patterns from the non-native group. These two groups with native speakers preferred QNonF compared to NF for both speaker nationalities, but the difference was more pronounced for American speakers. Non-native speakers preferred NF compared to QNonF for both Dutch and American speakers. This suggests that native English listeners showed the opposite pattern as non-native English listeners.

We can conclude that the speaker nationality influenced the decision of the listener (whether the same keyword sounded more native in NF or QNonF) for some keyword categories and that this differed between the two orders of the sound files. Listeners showed significantly different patterns depending on the order of the sound files, namely that they preferred the second sound file compared to the first one. Native listeners preferred QNonF compared to NF for both speaker nationalities, whereas non-native listeners preferred the opposite.

6 References

Baayen, R.H. (2013). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics".. R package version 1.4.1. https://CRAN.Rproject.org/package=languageR

Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from http://www.praat.org

CMU pronouncing dictionary, Carnegie Mellon University, Pittsburgh, PA, USA, http://www.speech.cs.cmu.edu/cgi-bin/cmudict, accessed by Katherine Marcoux in 2018

Cooke, M., & Lecumberri, M. L. (2012). The intelligibility of Lombard speech for non-native listeners. *The Journal of the Acoustical Society of America, 132*(2), 1120-1129.

Davies, A. (2015). Foreign accent: The phenomenon of non-native speech. *Journal of Multilingual and Multicultural Development, 36*(5), 545-548. doi:10.1080/01434632.2014.973284

Dreher, J., & O'Neill, J. (1957). Effects of ambient noise on speaker intelligibility for words and phrases, *The Journal of the Acoustical Society of America, 29,* 1320–1323.

Eurostat, European Union, Statistics Explained, Foreign language learning statistics, https://ec.europa.eu/eurostat/statistics-explained/index.php/Foreign language_skills_statistics#Number_of_foreign_languages_known, accessed on January 27, 2019

Hanulíková, A., & Weber, A. (2011). Sink positive: Linguistic experience with th substitutions influences non-native word recognition. *Attention, Perception, and Psychophysics, 74*(3), 613-629. doi:10.3758/s13414-011-0259-7

Jesney, K. (2004). The use of global foreign accent rating in studies of L2 acquisition. Unpublished Master's thesis, University of Calgary, Language Research Center.

Junqua, J. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers. *The Journal of the Acoustical Society of America, 93*(1), 510-524. doi:10.1121/1.405631

Langdon, H. W. (1999). Foreign Accent: Implications for Delivery of Speech and Language service. *Topics in Language Disorders, 19*(4), 49-65. doi:10.1097/00011363-199908000-00008

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

Li, C. (2003). Accent, intelligibility, and comprehensibility in the perception of foreign-accented Lombard speech. *The Journal of the Acoustical Society of America, 114*, 2364 doi:10.1121/1.4809202

LimeSurvey GmbH / LimeSurvey: An Open Source survey tool /LimeSurvey GmbH, Hamburg, Germany. URL http://www.limesurvey.org

Lombard, E. (1911). Le signe d'e´le´vation de la voix (The sign of the elevation of the voice), *Ann. Maladies l'oreille Larynx, 37*, 101–119. (translated by Paul H. Mason 2006, reviewed by M. Jacques Lenoir)

Lu, Y., & Cooke, M. (2009). Speech production modifications produced in the presence of low-pass and high-pass filtered noise. *The Journal of the Acoustical Society of America, 126*(3), 1495-1499. doi:10.1121/1.3179668

Lu, Y., & Cooke, M. (2008). Speech production modifications produced by competing talkers, babble, and stationary noise. *The Journal of the Acoustical Society of America, 124*(5), 3261-3275. doi:10.1121/1.2990705

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology, 59*(3), 203-243. doi:10.1016/j.cogpsych.2009.04.001

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th Conference of the International Speech Communication Association*.

Pisoni, D., Bernacki, R., Nusbaum, H., & Yuchtman, M. (1985). Some acoustic-phonetic correlates of speech produced in noise. *ICASSP 85. IEEE International Conference on Acoustics, Speech, and Signal Processing*. *10*,1581-1584. doi: 10.1109/icassp.1985.1168217

Pittman, A. L., & Wiley, T. L. (2001). Recognition of Speech Produced in Noise. *Journal of Speech Language and Hearing Research, 44*(3), 487. doi:10.1044/1092-4388 (2001/038)

Qualtrics software, Copyright © [2019] Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. https://www.Qualtrics.com

R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Romaine, S. (1996). Bilingualism. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 571-604). San Diego, CA: Academic Press.

Simon, E. (2010). Phonological transfer of voicing and devoicing rules: Evidence from L1 Dutch and L2 English conversational speech. *Language Sciences, 32*(1), 63-86. doi:10.1016/j.langsci.2008.10.001

Upton, G., & Cook, I. (2014). *A dictionary of statistics* (Third ed., Oxford paperback reference) [Third edition.]. Oxford: Oxford University Press. (2014). Retrieved February, 2019, from http://www.oxfordreference.com/view/10.1093/oi/authority.20110803100258143

van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analysis. *The Journal of the Acoustical Society of America, 84*, 917–928, doi: 10.1121/1.396660

Vokic, G. (2010). L1 allophones in L2 speech production: The case of English learners of Spanish. *Hispania, 93*(3), 430-452.

WebExp, Webexperimenten van de Radboud Universiteit, https://www.ru.nl/letteren/webexp/webexperimenten/, accessed in November and December 2018

7 Appendices

7.1 Experiment 1

7.1.1 Accentedness Ratings of Speakers

Experiment 1a	Participant	Experiment 1b	Participant	Experiment 1b - Experiment 1a
2,44	DU18	2,30	DU18	-0,15
3,58	DU27	2,81	DU26	-0,77
3,64	DU04	2,83	DU27	-0,81
3,89	DU21	2,89	DU28	-1,00
3,92	DU28	3,04	DU21	-0,88
3,94	DU32	3,11	DU32	-0,83
4,31	DU26	3,19	DU04	-1,12
4,33	DU35	3,26	DU25	-1,07
4,44	DU02	3,59	DU02	-0,85
4,47	DU25	3,67	DU37	-0,81
4,56	DU01	3,94	DU12	-0,61
4,69	DU09	3,96	DU09	-0,73
4,78	DU11	4,00	DU01	-0,78
5,00	DU31	4,09	DU11	-0,91
5,03	DU24	4,09	DU35	-0,94
5,03	DU10	4,11	DU10	-0,92
5,08	DU37	4,20	DU38	-0,88
5,19	DU12	4,26	DU24	-0,94
5,31	DU38	4,31	DU07	-0,99
5,64	DU33	4,31	DU33	-1,32
5,86	DU06	4,46	DU06	-1,40
5,89	DU07	4,59	DU31	-1,30
5,94	DU03	4,65	DU03	-1,30

Selected speakers = blue or purple, speakers that were selected in both studies = blue, speakers that were selected once = purple, speakers with very high scores in both pilots (strong foreign accent) = red, speaker with very low scores in both pilots (weak foreign accent) = green

7.1.2 Participants

Pilot	Age	gender	L1: English	Monolingual upbringing	Knowledge of Dutch	Knowledge of German	Average across all trials
1	26	female	yes	yes	no	no	5.03
1	25	female	yes	yes	no	no	3.76
1	28	female	yes	yes	no	no	3.87
1	24	female	yes	yes	no	no	4.33
1	24	male	yes	yes	no	no	4.45
1	26	female	yes	yes	no	no	4.73
2	32	male	yes	yes	no	no data	3.57
2	26	male	yes	yes	no	no data	2.24
2	28	male	yes	yes	no	no data	3.72
2	19	female	yes	yes	no	no data	3.44

2	26	male	yes	yes	no	no data	3.36
2	24	other	yes	yes	no	no data	4.30
2	26	male	yes	yes	no	no data	4.48
2	25	male	yes	yes	no	no data	3.25
2	30	female	yes	yes	no	no data	3.36

7.2 Experiment 2

7.2.1 Keywords

words with initial theta:	words with final voiced obstruents:	words with schwa in pre-stress position:
theater	blood	balloon
theology	club	banana
theory	crib	botanical
therapist	lemonade	cadaver
thermal	neighborhood	computer
thermos	rehab	gorilla
theta	road	parade
thriller	wood	police
throne		professor
		salami
		tomato

7.2.2 Carrier Sentences of Keywords

balloon

Did Sarah play with a **butterfly** all morning? No, she played with a **balloon** all morning.

Did **James** play with a balloon all morning? No, **Sarah** played with a balloon all morning.

Did Amanda hold a **kitten** all day? No, she held a **balloon** all day.

Did **Jessica** hold the balloon all day? No, **Amanda** held the balloon all day.

cadaver

Did the doctor examine the **patient** in detail? No, he examined the **cadaver** in detail.

Did the **assistant** examine the cadaver in detail? No, the **doctor** examined the cadaver in detail.

Did the examiner place the **liver** on the table? No, he placed the **cadaver** on the table.

Did the **police** place the cadaver on the table? No, the **examiner** placed the cadaver on the table.

computer

Did your brother buy a new **television** online? No, he bought a new **computer** online.

Did your **sister** buy a new computer online? No, my **brother** bought a new computer online.

Did your friend fix the **freezer** in five minutes? No, she fixed the **computer** in five minutes.

Did your **neighbor** fix the computer in five minutes? No, my **friend** fixed the computer in five minutes.

gorilla

Did Elizabeth see a **chimpanzee** in the zoo? No, she saw a **gorilla** in the zoo.

Did **Katherine** see a gorilla in the zoo? No, **Elisabeth** saw a gorilla in the zoo.

Did the woman photograph an **elephan**t in the wild? No, she photographed a **gorilla** in the wild.

Did the **funny** woman photograph a gorilla in the wild? No, the **serious** woman photographed a gorilla in the wild.

banana

Did Sophie eat an **apple** with her lunch today? No, she ate a **banana** with her lunch today.

Did Anna eat a **banana** with her lunch today? No, Sophie ate a **banana** with her lunch today.

Did Robert buy a **yogurt** for his snack later? No, he bought a **banana** for his snack later.

Did **Chris** buy a yogurt for his snack later? No, **Robert** bought a banana for his snack later.

professor

Did Diana see her **sister** on a walk on the beach? No, she saw her **professor** on a walk on the beach.

Did **Rose** see her professor on a walk on the beach? No, **Diana** saw her professor on a walk on the beach.

Did Paul meet his **friend** at the café to talk? No, he met his **professor** at the café to talk.

Did **Simon** meet his professor at the café to talk? No, **Paul** met his professor at the café to talk.

tomato

Will his father check whether the **orange** is ripe? No, he will check whether the **tomato** is ripe.

Will his **mother** check whether the tomato is ripe? No, his **father** will check whether the tomato is ripe.

Did the child ask if the **apple** was sweet? No, she asked if the **tomato** was sweet.

Did the **man** ask if the tomato was sweet? No, the **child** asked if the tomato was sweet.

botanical

Did the children visit the **butterfly** garden in town? No, they visited the **botanical** garden in town.

Did the **grandparents** visit the botanical garden in town? No, the **children** visited the botanical garden in town.

Did your friends explore the **Japanese** garden in the city? No, they explored the **botanical** garden in the city.

Did your **cousins** explore the botanical garden in the city? No, my **friends** explored the botanical garden in the city.

salami

Does Will like **chicken** in his sandwiches for lunch? No, he likes **salami** in his sandwiches for lunch.

Does **Matt** like salami in his sandwiches for lunch? No, **Will** likes salami in his sandwiches for lunch.

Did Claire want **peppers** on her pizza for dinner? No, she wanted **salami** on her pizza for dinner.

Did **Lisa** want salami on her pizza for dinner? No, **Claire** wanted salami on her pizza for dinner.

parade

Did the family go to the **beach** in Barcelona? No, they went to the **parade** in Barcelona.

Did the **friends** go to the parade in Barcelona? No, the **family** went to the parade in Barcelona.

Did Lily enjoy the flower **garden** in the spring? No, she enjoyed the flower **parade** in the spring.

Did **Ellen** enjoy the flower parade in the spring? No, **Lily** enjoyed the flower parade in the spring.

theater

Does James enjoy spending time in a **zoo** on Sundays? No, he enjoys spending time in a **theater** on Sundays.

Does **Henry** enjoy spending time in a theater on Sundays? No, **James** enjoys spending time in a theater on Sundays.

Does Betty like working in a **museum** in Paris? No, she likes working in a **theater** in Paris.

Does **Carmen** like working in a theater in Paris? No, **Betty** likes working in a theater in Paris.

therapist

Does Emma visit her **dentist** every week? No, she visits her **therapist** every week.

Does **Alice** visit her therapist every week? No, **Emma** visits her therapist every week.

Does Erin see her **tutor** every month? No, she sees her **therapist** every month.

Does **Heather** see her therapist every month? No, **Erin** sees her therapist every month.

theology

Will the students discuss general **philosophy** in modern day? No, they will discuss general **theology** in modern day.

Will the parents discuss general **theology** in modern day? No, the students will discuss general **theology** in modern day.

Will the lunch talk focus on **linguistics** in higher education? No, it will focus on **theology** in higher education.

Will the **morning** talk focus on theology in higher education? No, the **lunch** talk will focus on theology in higher education.

theory

Did Mary read about the general **implications** of relativity? No, she read about the general **theory** of relativity.

Did **Sally** read about the general theory of relativity? No, **Mary** read about the general theory of relativity.

Did Dan believe the clever **lie** Steven told him? No, he believed the clever **theory** Steven told him.

Did **Brian** believe the clever theory Steven told him? No, **Dan** believed the clever theory Steven told him.

thermal

Will Adam buy **waterproof** gloves for skiing? No, he will buy **thermal** gloves for skiing.

Will **Matt** buy thermal gloves for skiing? No, **Adam** will buy thermal gloves for skiing.

Will Rose wear **summer** pants for hiking? No, she will wear **thermal** pants for hiking.

Will **Mary** wear thermal pants for hiking? No, **Rose** will wear thermal pants for hiking.

thermos

Did Alan leave his gray **bottle** on his desk? No, he left his gray **thermos** on his desk.

Did **Nick** leave his gray thermos on his desk? No, **Alan** left his gray thermos on his desk.

Did Katie place her blue **bottle** on the table? No, she placed her blue **thermos** on the table.

Did **Daisy** place her blue thermos on the table? No, **Katie** placed her blue thermos on the table.

thriller

Will Julie read a new **comedy** over the weekend? No, she will read a new **thriller** over the weekend.

Will **Lizzy** read a new thriller over the weekend? No, **Julie** will read a new thriller over the weekend.

Will your sister watch a **drama** tomorrow? No, she will watch a **thriller** tomorrow.

Will your **oldest** sister watch a thriller tomorrow? No, my **youngest** sister will watch a thriller tomorrow.

throne

Will the queen sit on that comfy **chair** if she has guests? No, she will sit on that comfy **throne** if she has guests.

Will the **princess** sit on that comfy throne if she has guests? No, the **queen** will sit on that comfy throne if she has guests.

Will the boy clean the empty **staircase** if he has time? No, he will clean the empty **throne** if he has time.

Will the **tall** boy clean the empty throne if he has time? No, the short boy will clean the empty throne if he has time.

pub

Did the man walk into the **restaurant** in the city? No, he walked into the **pub** in the city.

Did the **fat** man walk into the pub in the city? No, the **thin** man walked into the pub in the city.

Did the woman drive to the **theater** in town? No, she drove to the **pub** in town.

Did the **short** woman drive to the pub in town? No, the **tall** woman drove to the pub in town.

cab

Did the man say the **bike** was yellow? No, he said the **cab** was yellow.

Did the **woman** say the cab was yellow? No, the **man** said the cab was yellow.

Did the lady ask if the **bus** was red? No, she asked if the **cab** was red.

Did the **child** ask if the cab was red? No, the **lady** asked if the cab was red.

club

Do your parents listen to jazz at the **café** in town? No, they listen to jazz at the **club** in town.

Do your **cousins** listen to jazz at the club in town? No, my **parents** listen to jazz at the club in town.

Do your friends dance in the **park** on Fridays? No, they dance in the **club** on Fridays.

Do your **brothers** dance in the club on Fridays? No, my **friends** dance in the club on Fridays.

lab

Will your neighbor work in the **library** across the street? No, he will work in the **lab** across the street.

Will your **kind** neighbor work in the lab across the street? No, my **mean** neighbor will work in the lab across the street.

Will your brother go to the **pharmacy** across the square? No, he will go to the **lab** across the square.

Will your **older** brother go to the lab across the square? No, my **younger** brother will go to the lab across the square.

food

Will your sister serve French **wine** on her birthday? No, she will serve French **food** on her birthday.

Will your **younger** sister serve French food on her birthday? No, my **older** sister will serve French food on her birthday.

Will David have Spanish **music** at the party on Saturday? No, he will have Spanish **food** at the party on Saturday.

Will **Alex** have Spanish food at the party on Saturday? No, **David** will have Spanish food at the party on Saturday.

crib

Did the baby sleep in his own **bed** all night? No, he slept in his own **crib** all night.

Did the **youngest** baby sleep in his own crib all night? No, the **oldest** baby slept in his own crib all night.

Did the parents build the **table** in one day? No, they built the **crib** in one day.

Did the **mother** build the crib in one day? No, the **parents** built the crib in one day.

rehab

Will the celebrity go to **jail** in the US? No, the celebrity will go to **rehab** in the US.
Will the **politician** go to rehab in the US? No, the **celebrity** will go to rehab in the US.
Will Roy work in **prison** as a counselor next year? No, he will work in **rehab** as a counselor next year.
Will **Bob** work in rehab as a counselor next year? No, **Roy** will work in rehab as a counselor next year.

blood

Did the laboratory store **urine** in those tubes? No, it stored **blood** in those tubes.
Did the **hospital** store blood in those tubes? No, the **laboratory** stored blood in those tubes.
Did the assistant take **saliva** as a sample? No, he took **blood** as a sample.
Did the **doctor** take blood as a sample? No, the **assistant** took blood as a sample.

wood

Did John buy **matches** at the store for a bonfire? No, he bought **wood** at the store for a bonfire.
Did **Paul** buy wood at the store for a bonfire? No, **John** bought wood at the store for a bonfire.
Did Simon use **cement** as a foundation for the house? No, he used **wood** as a foundation for the house.
Did **Max** use wood as a foundation for the house? No, **Simon** used wood as a foundation for the house.

road

Did the chicken cross the **garden** in the rain? No, it crossed the **road** in the rain.
Did the **turkey** cross the road in the rain? No, the **chicken** crossed the road in the rain.
Did the girl walk across the **park** in the sun? No, she walked across the **road** in the sun.
Did the **boy** walk across the road in the sun? No, the **girl** walked across the road in the sun.

lemonade

Will the kids drink **water** in the summer? No, they will drink **lemonade** in the summer.
Will the **toddlers** drink lemonade in the summer? No, the **kids** will drink lemonade in the summer.
Will your cousins make **tea** in July? No, they will make **lemonade** in July.
Will your **neighbors** make lemonade in July? No, my **cousins** will make lemonade in July.

neighborhood

Will your friend live on the best **street** in Dallas? No, she will live in the best **neighborhood** in Dallas.
Will your **cousin** live in the best neighborhood in Dallas? No, my **friend** will live in the best neighborhood in Dallas.
Will your sister move to a posh **house** in Miami? No, she will move to a posh **neighborhood** in Miami.
Will your **brother** move to a posh neighborhood in Miami? No, my **sister** will move to a posh neighborhood in Miami.

7.2.3 Demographic Questions

1. How did you choose the word that sounded more native to you?

2. Do you have any comments on the experiment?

3. How old are you? ____ years
4. What is your gender?
    O female
    O male
    O other
5. What is the highest level of education you have completed?
    O Some high school, no degree
    O High school
    O Some college, no degree
    O Vocational training
    O Bachelor's degree
    O Master's degree
    O Professional degree
    O Doctorate degree
    O other: _____
6. Is English your native language?
    O yes
    O no
7. Did you grow up multilingually (speaking two or more languages)?
    O yes, I spoke these languages: _____
    O no
8. What is/are your parents' native language(s)?

9. Have you ever lived abroad? If so, in which country/countries did you live and for how many months?
    O yes: _____
    O no
10. Which foreign languages do you speak and how proficient are you in those? Please indicate your proficiency level in each language by writing "beginner," "intermediate," "fluent," or "native-like" next to the language.

11. What is your level of proficiency in Dutch?
    O none
    O beginner
    O intermediate
    O high
    O native-like
12. How many hours per week do you speak English with non-native English speakers? This includes conversations at work/at school and with friends.

13. What is/are the native language(s) of these non-native speakers of English you talk to regularly?

7.2.4 Foreign Languages Spoken by the Participants

	final devoicing	schwa	usable
Spanish	no	no	no
French	no	yes	yes
Italian	no	no	no
Portuguese	maybe	no	no
German	yes	yes	no
Dutch	yes	yes	no
Danish	no	yes	yes
Serbio-Croation	no information	no	no
Bosnian	no information	no	no
Macedonian	yes	yes	no
Tagalog/Filipino	no	no	no
Russian	yes	yes	no
Latvian	no	no	no
Korean	yes	no	no
Japanese	no	no	no
Mandarin	no	yes	yes
Cantonese	no	no	no
Hokkien/Teochew	no	yes	yes
Hakka	no	no	no
Gujarati	no	yes	yes
Punjabi	no	yes	yes
Hindi	no	yes	yes
Crow	no	yes, in diphthongs	yes
Arabic	no	yes	yes
Shona	no	no	no
Kurdish	no	no	no
Rwandan	no info	no	no

7.2.5 Experimental Lists

L01=[DU10, DU11, EN01, EN04, DU025, DU024, DU01, EN09, EN02, DU031, EN03, EN07, EN08, DU9, DU02, EN06]
L02=[DU031, EN02, EN07, EN03, DU02, DU11, EN01, DU10, EN09, DU9, EN06, DU01, EN04, EN08, DU024, DU025]
L03=[DU9, EN01, EN04, DU024, DU11, EN03, DU02, DU031, EN09, EN06, EN07, DU10, EN02, DU025, EN08, DU01]
L04=[DU01, DU024, DU10, EN08, EN01, EN07, DU031, EN02, DU025, EN03, DU02, DU9, EN04, EN09, DU11, EN06]
L05=[DU10, DU31, EN09, DU02, EN04, EN08, DU25, DU11, DU09, EN07, EN06, DU01, EN02, DU24, EN03, EN01]
L06=[DU24, EN03, DU25, DU02, EN07, DU09, DU11, EN08, DU31, EN02, DU01, EN01, DU10, EN06, EN04, EN09]
L07=[EN04, EN06, DU01, DU10, EN08, EN07, EN01, DU031, DU02, EN09, DU9, DU024, DU025, EN03, EN02, DU11]

L08=[EN09, DU02, DU024, DU9, EN02, EN06, DU01, EN04, DU031, EN03, DU11, EN01, DU10, DU025, EN07, EN08]
L09=[EN03, DU01, DU10, EN07, EN06, DU9, EN02, EN09, DU031, DU11, DU024, EN04, DU02, EN08, DU025, EN01]
L10=[EN01, EN07, EN04, DU025, DU01, DU024, EN09, DU02, EN08, DU9, EN02, EN03, DU10, DU031, EN06, DU11]
L11=[EN04, EN09, DU031, EN02, DU10, DU025, EN08, EN06, EN03, DU024, DU11, EN01, DU02, EN07, DU9, DU01]
L12=[EN07, DU9, EN08, EN02, DU024, EN03, EN06, DU025, EN09, DU02, EN01, DU01, EN04, DU11, DU10, DU031]

7.2.6 Example of Experimental List

list1a=[DU10, DU11, EN01, EN04, DU025, DU024, DU01, EN09, EN02, DU031, EN03, EN07, EN08, DU9, DU02, EN06]

list1a=[['DU10_thermos_o2.wav', 'DU10_parade_o1.wav', 'DU10_road_o2.wav', 'DU10_throne_o2.wav', 'DU10_theta_o2.wav', 'DU10_banana_o1.wav', 'DU10_theory_o1.wav', 'DU10_rehab_o1.wav', 'DU10_therapist_o1.wav', 'DU10_wood_o1.wav', 'DU10_thriller_o2.wav', 'DU10_balloon_o1.wav', 'DU10_salami_o1.wav', 'DU10_club_o1.wav', 'DU10_blood_o2.wav', 'DU10_cadaver_o2.wav', 'DU10_theater_o1.wav', 'DU10_professor_o1.wav', 'DU10_lemonade_o2.wav', 'DU10_computer_o1.wav', 'DU10_crib_o2.wav', 'DU10_neighborhood_o1.wav', 'DU10_botanical_o2.wav', 'DU10_tomato_o2.wav', 'DU10_police_o2.wav', 'DU10_thermal_o2.wav', 'DU10_gorilla_o2.wav'], ['DU11_therapist_o1.wav', 'DU11_rehab_o2.wav', 'DU11_gorilla_o2.wav', 'DU11_throne_o2.wav', 'DU11_tomato_o2.wav', 'DU11_salami_o1.wav', 'DU11_computer_o1.wav', 'DU11_theta_o2.wav', 'DU11_botanical_o2.wav', 'DU11_professor_o2.wav', 'DU11_club_o2.wav', 'DU11_thriller_o1.wav', 'DU11_wood_o1.wav', 'DU11_theater_o1.wav', 'DU11_road_o1.wav', 'DU11_thermal_o2.wav', 'DU11_neighborhood_o1.wav', 'DU11_lemonade_o1.wav', 'DU11_banana_o2.wav', 'DU11_crib_o2.wav', 'DU11_theory_o2.wav', 'DU11_thermos_o1.wav', 'DU11_parade_o1.wav', 'DU11_theology_o2.wav', 'DU11_balloon_o2.wav', 'DU11_cadaver_o1.wav', 'DU11_police_o1.wav', 'DU11_blood_o1.wav'], ['EN01_theory_o2.wav', 'EN01_thermos_o1.wav', 'EN01_theta_o2.wav', 'EN01_crib_o2.wav', 'EN01_balloon_o1.wav', 'EN01_cadaver_o1.wav', 'EN01_club_o1.wav', 'EN01_botanical_o1.wav', 'EN01_neighborhood_o1.wav', 'EN01_theology_o1.wav', 'EN01_thriller_o1.wav', 'EN01_professor_o2.wav', 'EN01_gorilla_o2.wav', 'EN01_therapist_o2.wav', 'EN01_salami_o1.wav', 'EN01_computer_o1.wav', 'EN01_tomato_o2.wav', 'EN01_wood_o2.wav', 'EN01_banana_o2.wav', 'EN01_road_o1.wav', 'EN01_parade_o2.wav', 'EN01_police_o1.wav', 'EN01_blood_o2.wav', 'EN01_rehab_o2.wav', 'EN01_lemonade_o2.wav', 'EN01_throne_o1.wav', 'EN01_theater_o2.wav', 'EN01_thermal_o1.wav'], ['EN04_cadaver_o2.wav', 'EN04_computer_o1.wav', 'EN04_banana_o1.wav', 'EN04_rehab_o2.wav', 'EN04_theory_o2.wav', 'EN04_police_o1.wav', 'EN04_gorilla_o2.wav', 'EN04_parade_o2.wav', 'EN04_lemonade_o2.wav', 'EN04_professor_o2.wav', 'EN04_thermos_o2.wav', 'EN04_botanical_o1.wav', 'EN04_salami_o1.wav', 'EN04_thermal_o2.wav', 'EN04_thriller_o2.wav', 'EN04_crib_o2.wav', 'EN04_therapist_o2.wav', 'EN04_club_o2.wav', 'EN04_theater_o1.wav', 'EN04_theta_o2.wav', 'EN04_balloon_o1.wav', 'EN04_theology_o1.wav', 'EN04_tomato_o1.wav', 'EN04_blood_o1.wav', 'EN04_throne_o1.wav', 'EN04_road_o1.wav', 'EN04_neighborhood_o1.wav', 'EN04_wood_o1.wav'], ['DU25_cadaver_o1.wav', 'DU25_gorilla_o1.wav', 'DU25_theater_o2.wav', 'DU25_theory_o1.wav', 'DU25_therapist_o1.wav', 'DU25_road_o2.wav', 'DU25_thermos_o2.wav', 'DU25_theology_o2.wav', 'DU25_throne_o2.wav', 'DU25_professor_o1.wav', 'DU25_salami_o1.wav', 'DU25_neighborhood_o1.wav', 'DU25_lemonade_o2.wav', 'DU25_balloon_o1.wav', 'DU25_thriller_o2.wav', 'DU25_crib_o2.wav', 'DU25_thermal_o2.wav', 'DU25_computer_o1.wav', 'DU25_banana_o2.wav', 'DU25_rehab_o1.wav', 'DU25_club_o2.wav', 'DU25_wood_o2.wav', 'DU25_theta_o2.wav', 'DU25_tomato_o1.wav', 'DU25_blood_o1.wav', 'DU25_parade_o1.wav', 'DU25_police_o1.wav'], ['DU24_computer_o2.wav', 'DU24_theater_o1.wav', 'DU24_balloon_o2.wav', 'DU24_professor_o2.wav', 'DU24_thermal_o1.wav', 'DU24_neighborhood_o1.wav', 'DU24_thermos_o2.wav', 'DU24_lemonade_o2.wav', 'DU24_blood_o2.wav', 'DU24_botanical_o2.wav', 'DU24_theology_o2.wav', 'DU24_crib_o2.wav', 'DU24_tomato_o1.wav', 'DU24_gorilla_o1.wav', 'DU24_wood_o2.wav', 'DU24_rehab_o1.wav', 'DU24_parade_o1.wav', 'DU24_theory_o1.wav', 'DU24_throne_o1.wav', 'DU24_banana_o1.wav', 'DU24_cadaver_o1.wav', 'DU24_thriller_o1.wav', 'DU24_club_o2.wav', 'DU24_road_o2.wav', 'DU24_therapist_o1.wav', 'DU24_salami_o2.wav', 'DU24_police_o1.wav'], ['DU01_police_o2.wav', 'DU01_salami_o2.wav', 'DU01_tomato_o2.wav', 'DU01_road_o2.wav', 'DU01_gorilla_o1.wav', 'DU01_therapist_o1.wav', 'DU01_lemonade_o1.wav', 'DU01_theta_o2.wav', 'DU01_thermal_o2.wav', 'DU01_blood_o1.wav', 'DU01_wood_o2.wav', 'DU01_thermos_o1.wav', 'DU01_rehab_o2.wav', 'DU01_parade_o2.wav', 'DU01_theology_o1.wav', 'DU01_botanical_o1.wav', 'DU01_balloon_o2.wav', 'DU01_professor_o1.wav', 'DU01_theater_o2.wav', 'DU01_computer_o1.wav', 'DU01_club_o2.wav', 'DU01_banana_o1.wav', 'DU01_cadaver_o1.wav', 'DU01_crib_o1.wav', 'DU01_neighborhood_o2.wav', 'DU01_theory_o2.wav', 'DU01_throne_o1.wav'], ['EN09_thermal_o1.wav', 'EN09_wood_o2.wav', 'EN09_rehab_o1.wav', 'EN09_cadaver_o2.wav', 'EN09_tomato_o1.wav', 'EN09_balloon_o1.wav', 'EN09_neighborhood_o2.wav', 'EN09_road_o2.wav', 'EN09_theta_o2.wav', 'EN09_professor_o1.wav',

'EN09_lemonade_o1.wav', 'EN09_botanical_o1.wav', 'EN09_throne_o1.wav', 'EN09_parade_o2.wav', 'EN09_club_o2.wav', 'EN09_thriller_o2.wav', 'EN09_gorilla_o2.wav', 'EN09_theater_o1.wav', 'EN09_theology_o1.wav', 'EN09_blood_o1.wav', 'EN09_salami_o1.wav', 'EN09_therapist_o1.wav', 'EN09_police_o2.wav', 'EN09_thermos_o2.wav', 'EN09_banana_o2.wav', 'EN09_computer_o2.wav', 'EN09_theory_o1.wav', 'EN09_crib_o2.wav'], ['EN02_neighborhood_o1.wav', 'EN02_botanical_o1.wav', 'EN02_therapist_o2.wav', 'EN02_throne_o2.wav', 'EN02_salami_o1.wav', 'EN02_tomato_o2.wav', 'EN02_thermos_o1.wav', 'EN02_police_o1.wav', 'EN02_parade_o2.wav', 'EN02_professor_o1.wav', 'EN02_crib_o1.wav', 'EN02_theory_o1.wav', 'EN02_theology_o2.wav', 'EN02_theta_o1.wav', 'EN02_cadaver_o1.wav', 'EN02_road_o2.wav', 'EN02_thriller_o2.wav', 'EN02_balloon_o1.wav', 'EN02_rehab_o1.wav', 'EN02_banana_o2.wav', 'EN02_thermal_o2.wav', 'EN02_gorilla_o2.wav', 'EN02_lemonade_o2.wav', 'EN02_wood_o2.wav', 'EN02_theater_o2.wav', 'EN02_computer_o1.wav', 'EN02_blood_o2.wav', 'EN02_club_o1.wav'], ['DU31_computer_o1.wav', 'DU31_cadaver_o2.wav', 'DU31_neighborhood_o2.wav', 'DU31_banana_o2.wav', 'DU31_thermos_o1.wav', 'DU31_tomato_o1.wav', 'DU31_professor_o2.wav', 'DU31_lemonade_o2.wav', 'DU31_road_o1.wav', 'DU31_wood_o2.wav', 'DU31_balloon_o1.wav', 'DU31_police_o1.wav', 'DU31_thriller_o1.wav', 'DU31_botanical_o2.wav', 'DU31_theta_o1.wav', 'DU31_theology_o2.wav', 'DU31_parade_o2.wav', 'DU31_gorilla_o2.wav', 'DU31_club_o1.wav', 'DU31_throne_o1.wav', 'DU31_theater_o1.wav', 'DU31_salami_o2.wav', 'DU31_rehab_o1.wav', 'DU31_therapist_o2.wav', 'DU31_crib_o1.wav', 'DU31_theory_o1.wav', 'DU31_blood_o2.wav'], ['EN03_botanical_o2.wav', 'EN03_theater_o1.wav', 'EN03_thermos_o1.wav', 'EN03_salami_o1.wav', 'EN03_theta_o1.wav', 'EN03_wood_o1.wav', 'EN03_blood_o2.wav', 'EN03_computer_o2.wav', 'EN03_throne_o2.wav', 'EN03_rehab_o2.wav', 'EN03_theory_o1.wav', 'EN03_balloon_o1.wav', 'EN03_parade_o2.wav', 'EN03_therapist_o1.wav', 'EN03_tomato_o1.wav', 'EN03_gorilla_o2.wav', 'EN03_police_o2.wav', 'EN03_crib_o1.wav', 'EN03_theology_o1.wav', 'EN03_cadaver_o1.wav', 'EN03_neighborhood_o2.wav', 'EN03_lemonade_o1.wav', 'EN03_thriller_o2.wav', 'EN03_banana_o2.wav', 'EN03_professor_o1.wav', 'EN03_road_o2.wav', 'EN03_thermal_o2.wav', 'EN03_club_o2.wav'], ['EN07_lemonade_o2.wav', 'EN07_theater_o1.wav', 'EN07_police_o1.wav', 'EN07_botanical_o1.wav', 'EN07_blood_o2.wav', 'EN07_theology_o2.wav', 'EN07_thriller_o1.wav', 'EN07_theta_o1.wav', 'EN07_gorilla_o1.wav', 'EN07_therapist_o2.wav', 'EN07_wood_o2.wav', 'EN07_salami_o1.wav', 'EN07_parade_o2.wav', 'EN07_theory_o2.wav', 'EN07_balloon_o2.wav', 'EN07_computer_o1.wav', 'EN07_club_o2.wav', 'EN07_rehab_o1.wav', 'EN07_banana_o2.wav', 'EN07_tomato_o2.wav', 'EN07_thermal_o1.wav', 'EN07_road_o1.wav', 'EN07_thermos_o2.wav', 'EN07_cadaver_o2.wav', 'EN07_crib_o1.wav', 'EN07_professor_o1.wav', 'EN07_throne_o2.wav', 'EN07_neighborhood_o1.wav'], ['EN08_club_o2.wav', 'EN08_road_o2.wav', 'EN08_theta_o2.wav', 'EN08_throne_o2.wav', 'EN08_thriller_o1.wav', 'EN08_gorilla_o1.wav', 'EN08_blood_o1.wav', 'EN08_computer_o1.wav', 'EN08_botanical_o2.wav', 'EN08_cadaver_o2.wav', 'EN08_rehab_o1.wav', 'EN08_parade_o1.wav', 'EN08_therapist_o1.wav', 'EN08_balloon_o2.wav', 'EN08_police_o2.wav', 'EN08_neighborhood_o1.wav', 'EN08_professor_o2.wav', 'EN08_wood_o2.wav', 'EN08_theater_o2.wav', 'EN08_banana_o1.wav', 'EN08_lemonade_o1.wav', 'EN08_tomato_o1.wav', 'EN08_theology_o2.wav', 'EN08_theory_o2.wav', 'EN08_thermal_o1.wav', 'EN08_salami_o1.wav', 'EN08_crib_o2.wav', 'EN08_thermos_o1.wav'], ['DU09_balloon_o1.wav', 'DU09_club_o1.wav', 'DU09_banana_o2.wav', 'DU09_tomato_o1.wav', 'DU09_thermal_o1.wav', 'DU09_theology_o1.wav', 'DU09_blood_o1.wav', 'DU09_theory_o2.wav', 'DU09_professor_o1.wav', 'DU09_police_o1.wav', 'DU09_cadaver_o2.wav', 'DU09_lemonade_o2.wav', 'DU09_wood_o1.wav', 'DU09_road_o2.wav', 'DU09_computer_o2.wav', 'DU09_salami_o1.wav', 'DU09_crib_o2.wav', 'DU09_therapist_o2.wav', 'DU09_rehab_o2.wav', 'DU09_thriller_o1.wav', 'DU09_botanical_o2.wav', 'DU09_thermos_o1.wav', 'DU09_theta_o2.wav', 'DU09_throne_o1.wav', 'DU09_gorilla_o1.wav', 'DU09_theater_o2.wav', 'DU09_neighborhood_o2.wav', 'DU09_parade_o2.wav'], ['DU02_theta_o1.wav', 'DU02_balloon_o2.wav', 'DU02_blood_o1.wav', 'DU02_tomato_o1.wav', 'DU02_gorilla_o2.wav', 'DU02_salami_o2.wav', 'DU02_throne_o2.wav', 'DU02_crib_o1.wav', 'DU02_wood_o1.wav', 'DU02_theory_o2.wav', 'DU02_lemonade_o2.wav', 'DU02_therapist_o1.wav', 'DU02_thermal_o2.wav', 'DU02_club_o2.wav', 'DU02_parade_o2.wav', 'DU02_neighborhood_o1.wav', 'DU02_theater_o1.wav', 'DU02_computer_o2.wav', 'DU02_road_o1.wav', 'DU02_thriller_o2.wav', 'DU02_cadaver_o1.wav', 'DU02_theology_o1.wav', 'DU02_banana_o1.wav', 'DU02_rehab_o2.wav', 'DU02_professor_o1.wav', 'DU02_police_o2.wav', 'DU02_botanical_o2.wav', 'DU02_thermos_o1.wav'], ['EN06_therapist_o1.wav', 'EN06_throne_o1.wav', 'EN06_parade_o2.wav', 'EN06_balloon_o1.wav', 'EN06_theology_o2.wav', 'EN06_lemonade_o1.wav', 'EN06_thermal_o2.wav', 'EN06_theta_o2.wav', 'EN06_botanical_o2.wav', 'EN06_police_o2.wav', 'EN06_neighborhood_o1.wav', 'EN06_thriller_o1.wav', 'EN06_wood_o1.wav', 'EN06_theater_o2.wav', 'EN06_cadaver_o2.wav', 'EN06_computer_o1.wav', 'EN06_banana_o1.wav', 'EN06_rehab_o1.wav', 'EN06_blood_o1.wav', 'EN06_professor_o2.wav', 'EN06_crib_o2.wav', 'EN06_club_o2.wav', 'EN06_tomato_o1.wav', 'EN06_road_o2.wav', 'EN06_gorilla_o2.wav', 'EN06_salami_o1.wav', 'EN06_theory_o1.wav', 'EN06_thermos_o2.wav']]

7.2.7 WebExp Script

```
experiment {
	info {
	display="Welcome to the our experiment! We are researching how our recordings of English words are perceived in terms of nativeness. Every time you will hear two English words. Please indicate which one sounds more native to you. Please press Z if the first one sounds more native and M if the second one sounds more native. We will start with four practice trials."
	continue_text="Start"
		font {
			size=20
			color="black"
		}
	}
	lexicaldecision {
		decision {
		display="Which one sounds more native to you? Press Z (the first one) or M (the second one)."
		advancekeys="zZmM"
		sound="c_o1_girl_ppt0__mono.wav"
		}
		pauze {
			showtime=1
		}
		decision {
		display="Which one sounds more native to you? Press Z (the first one) or M (the second one)."
		advancekeys="zZmM"
		sound="c_o2_sister_ppt0__mono.wav"
		}
		pauze {
			showtime=1
		}
		decision {
		display="Which one sounds more native to you? Press Z (the first one) or M (the second one)."
		advancekeys="zZmM"
		sound="c_o2_job_ppt0__mono.wav"
		}
		pauze {
			showtime=1
		}
		decision {
		display="Which one sounds more native to you? Press Z (the first one) or M (the second one)."
		advancekeys="zZmM"
		sound="c_o1_parents_ppt0__mono.wav"
		}
	}
	info {
	display="Do you have any questions before the experiment starts? If not, click on the button to start the experiment."
	continue_text="Start"
		font {
			size=20
			color="black"
```

```
		}
	}
	include {
	files=["list01a_1.inc", "list01b_1.inc", "list02a_1.inc", "list02b_1.inc", "list03a_1.inc", "list03b_1.inc", "list04a_1.inc",
	"list04b_1.inc", "list05a_1.inc", "list05b_1.inc", "list06a_1.inc", "list06b_1.inc", "list07a_1.inc", "list07b_1.inc",
	"list08a_1.inc", "list08b_1.inc", "list09a_1.inc", "list09b_1.inc", "list10a_1.inc", "list10b_1.inc", "list11a_1.inc",
	"list11b_1.inc", "list12a_1.inc", "list12b_1.inc"]
	}
	info {
	display="This is the first break. Please click on the button, once you are ready to continue."
	continue_text="Continue"
		font {
			size=20
			color="black"
		}
	}
	include {
	files=["list01a_2.inc", "list01b_2.inc", "list02a_2.inc", "list02b_2.inc", "list03a_2.inc", "list03b_2.inc", "list04a_2.inc",
	"list04b_2.inc", "list05a_2.inc", "list05b_2.inc", "list06a_2.inc", "list06b_2.inc", "list07a_2.inc", "list07b_2.inc",
	"list08a_2.inc", "list08b_2.inc", "list09a_2.inc", "list09b_2.inc", "list10a_2.inc", "list10b_2.inc", "list11a_2.inc",
	"list11b_2.inc", "list12a_2.inc", "list12b_2.inc" ]
	}
	info {
		display="This is the second break. Please click on the button, once you are ready to continue."
		continue_text="Continue"
		font {
			size=20
			color="black"
		}
	}
	include {
	files=["list01a_3.inc", "list01b_3.inc", "list02a_3.inc", "list02b_3.inc", "list03a_3.inc", "list03b_3.inc", "list04a_3.inc",
	"list04b_3.inc", "list05a_3.inc", "list05b_3.inc", "list06a_3.inc", "list06b_3.inc", "list07a_3.inc", "list07b_3.inc",
	"list08a_3.inc", "list08b_3.inc", "list09a_3.inc", "list09b_3.inc", "list10a_3.inc", "list10b_3.inc", "list11a_3.inc",
	"list11b_3.inc", "list12a_3.inc", "list12b_3.inc" ]
	}
	info {
		display="This is the third break. Please click on the button, once you are ready to continue."
		continue_text="Continue"
		font {
			size=20
			color="black"
		}
	}
	include {
	files=["list01a_4.inc", "list01b_4.inc", "list02a_4.inc", "list02b_4.inc", "list03a_4.inc", "list03b_4.inc", "list04a_4.inc",
	"list04b_4.inc", "list05a_4.inc", "list05b_4.inc", "list06a_4.inc", "list06b_4.inc", "list07a_4.inc", "list07b_4.inc",
	"list08a_4.inc", "list08b_4.inc", "list09a_4.inc", "list09b_4.inc", "list10a_4.inc", "list10b_4.inc", "list11a_4.inc",
	"list11b_4.inc", "list12a_4.inc", "list12b_4.inc"]
	}
}
```

7.2.8 Excerpt from List1a Part 1

```
lexicaldecision {
        decision {
        display="Which one sounds more native? Press Z (first one) or M (second one)."
        advancekeys="zZmM"
        sound="DU10_thermos_o2.wav"
         }
        pauze {
        showtime=1
        }
        decision {
        display="Which one sounds more native? Press Z (first one) or M (second one)."
        advancekeys="zZmM"
        sound="DU10_parade_o1.wav"
         }
        pauze {
        showtime=1
        }
        decision {
        display="Which one sounds more native? Press Z (first one) or M (second one)."
        advancekeys="zZmM"
        sound="DU10_road_o2.wav"
         }
        pauze {
        showtime=1
        }
        decision {
        display="Which one sounds more native? Press Z (first one) or M (second one)."
        advancekeys="zZmM"
        sound="DU10_throne_o2.wav"
        }
        pauze {
        showtime=1
        }

        Etc.
```

7.2.9 Re-leveling of the Complete Model

Monolingual and multilingual plus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               0.419021   0.237364   1.765  0.07751 .
keyword_categoryschwa                    -0.100427   0.132244  -0.759  0.44761
keyword_categorytheta                    -0.018465   0.152129  -0.121  0.90339
speaker_nationalityEN                     0.015668   0.225306   0.070  0.94456
ordero2                                  -0.383247   0.096441  -3.974 7.07e-05 ***
listener_groupmonolingualminus            0.006357   0.274592   0.023  0.98153
listener_groupmultilingualminus          -0.370196   0.264289  -1.401  0.16130
listener_groupnon_native                 -0.809517   0.292867  -2.764  0.00571 **
keyword_categoryschwa:speaker_nationalityEN  0.428378   0.143715   2.981  0.00288 **
keyword_categorytheta:speaker_nationalityEN -0.308705   0.176765  -1.746  0.08074 .
keyword_categoryschwa:ordero2            -0.176843   0.071791  -2.463  0.01377 *
keyword_categorytheta:ordero2             0.132268   0.074942   1.765  0.07757 .
```

Monolingual and multilingual plus and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.40062    0.26683   1.501 0.133257
keyword_categoryfd                         0.01848    0.15218   0.121 0.903352
keyword_categoryschwa                     -0.08208    0.16267  -0.505 0.613844
speaker_nationalityEN                     -0.29306    0.27791  -1.055 0.291647
ordero2                                   -0.25096    0.09545  -2.629 0.008561 **
listener_groupmonolingualminus             0.00630    0.27476   0.023 0.981708
listener_groupmultilingualminus           -0.37025    0.26420  -1.401 0.161090
listener_groupnon_native                  -0.80946    0.29289  -2.764 0.005714 **
keyword_categoryfd:speaker_nationalityEN   0.30869    0.17687   1.745 0.080939 .
keyword_categoryschwa:speaker_nationalityEN  0.73723    0.20676   3.566 0.000363 ***
keyword_categoryfd:ordero2                -0.13226    0.07495  -1.765 0.077602 .
keyword_categoryschwa:ordero2             -0.30910    0.07037  -4.392 1.12e-05 ***
```

Monolingual and multilingual plus and schwa on intercept

(Failed to converge)

Multilingual minus and fd on intercept

```
Fixed effects:
                                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                                 0.04887    0.25464   0.192  0.84781
keyword_categoryschwa                      -0.10050    0.13220  -0.760  0.44714
keyword_categorytheta                      -0.01845    0.15205  -0.121  0.90343
speaker_nationalityEN                       0.01558    0.22524   0.069  0.94485
ordero2                                    -0.38320    0.09643  -3.974 7.07e-05 ***
listener_groupmonoandmultiplus              0.37026    0.26398   1.403  0.16073
listener_groupmonolingualminus              0.37655    0.30223   1.246  0.21280
listener_groupnon_native                   -0.43924    0.29628  -1.483  0.13820
keyword_categoryschwa:speaker_nationalityEN  0.42846    0.14368   2.982  0.00286 **
keyword_categorytheta:speaker_nationalityEN -0.30875    0.17669  -1.747  0.08057 .
keyword_categoryschwa:ordero2              -0.17686    0.07178  -2.464  0.01374 *
keyword_categorytheta:ordero2               0.13225    0.07493   1.765  0.07758 .
```

Multilingual minus and th on intercept

```
Fixed effects:
                                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                                 0.03050    0.27358   0.112 0.911220
keyword_categoryfd                          0.01844    0.15215   0.121 0.903523
keyword_categoryschwa                      -0.08208    0.16265  -0.505 0.613833
speaker_nationalityEN                      -0.29321    0.27779  -1.056 0.291194
ordero2                                    -0.25098    0.09546  -2.629 0.008560 **
listener_groupmonoandmultiplus              0.37012    0.26471   1.398 0.162051
listener_groupmonolingualminus              0.37653    0.30284   1.243 0.213742
listener_groupnon_native                   -0.43939    0.29688  -1.480 0.138865
keyword_categoryfd:speaker_nationalityEN    0.30874    0.17681   1.746 0.080786 .
keyword_categoryschwa:speaker_nationalityEN  0.73725    0.20671   3.567 0.000362 ***
keyword_categoryfd:ordero2                 -0.13225    0.07495  -1.765 0.077627 .
keyword_categoryschwa:ordero2              -0.30909    0.07038  -4.392 1.12e-05 ***
```

Multilingual minus and schwa on intercept

```
Fixed effects:
                                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                                -0.05159    0.26239  -0.197 0.844121
keyword_categoryfd                          0.10054    0.13218   0.761 0.446902
keyword_categorytheta                       0.08212    0.16256   0.505 0.613462
speaker_nationalityEN                       0.44405    0.25328   1.753 0.079561 .
ordero2                                    -0.56006    0.09293  -6.026 1.68e-09 ***
listener_groupmonoandmultiplus              0.37022    0.26442   1.400 0.161478
listener_groupmonolingualminus              0.37653    0.30268   1.244 0.213497
listener_groupnon_native                   -0.43930    0.29701  -1.479 0.139128
keyword_categoryfd:speaker_nationalityEN   -0.42854    0.14363  -2.984 0.002848 **
keyword_categorytheta:speaker_nationalityEN -0.73733    0.20660  -3.569 0.000359 ***
keyword_categoryfd:ordero2                  0.17684    0.07179   2.463 0.013762 *
keyword_categorytheta:ordero2               0.30910    0.07037   4.393 1.12e-05 ***
```

Non-native and fd on intercept

```
Fixed effects:
                                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                                -0.39036    0.26051  -1.498  0.13401
keyword_categoryschwa                      -0.10044    0.13224  -0.760  0.44750
keyword_categorytheta                      -0.01833    0.15217  -0.120  0.90412
speaker_nationalityEN                       0.01548    0.22539   0.069  0.94524
ordero2                                    -0.38319    0.09645  -3.973  7.1e-05 ***
listener_groupmonoandmultiplus              0.80949    0.29278   2.765  0.00569 **
listener_groupmonolingualminus              0.81585    0.35004   2.331  0.01977 *
listener_groupmultilingualminus             0.43916    0.29684   1.479  0.13902
keyword_categoryschwa:speaker_nationalityEN  0.42843    0.14369   2.982  0.00287 **
keyword_categorytheta:speaker_nationalityEN -0.30889    0.17684  -1.747  0.08069 .
keyword_categoryschwa:ordero2              -0.17687    0.07179  -2.464  0.01376 *
keyword_categorytheta:ordero2               0.13222    0.07494   1.764  0.07769 .
```

Non-native and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.40877    0.27452  -1.489 0.136486
keyword_categoryfd                         0.01846    0.15226   0.121 0.903512
keyword_categoryschwa                     -0.08209    0.16276  -0.504 0.614005
speaker_nationalityEN                     -0.29315    0.27798  -1.055 0.291617
ordero2                                   -0.25096    0.09546  -2.629 0.008564 **
listener_groupmonoandmultiplus             0.80937    0.29372   2.756 0.005859 **
listener_groupmonolingualminus             0.81575    0.35134   2.322 0.020243 *
listener_groupmultilingualminus            0.43914    0.29761   1.476 0.140072
keyword_categoryfd:speaker_nationalityEN   0.30874    0.17695   1.745 0.081021 .
keyword_categoryschwa:speaker_nationalityEN 0.73726   0.20684   3.564 0.000365 ***
keyword_categoryfd:ordero2                -0.13228    0.07495  -1.765 0.077588 .
keyword_categoryschwa:ordero2             -0.30912    0.07038  -4.392 1.12e-05 ***
```

Non-native and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.49090    0.26669  -1.841 0.065666 .
keyword_categoryfd                         0.10040    0.13227   0.759 0.447818
keyword_categorytheta                      0.08196    0.16269   0.504 0.614440
speaker_nationalityEN                      0.44389    0.25346   1.751 0.079886 .
ordero2                                   -0.56005    0.09293  -6.026 1.68e-09 ***
listener_groupmonoandmultiplus             0.80958    0.29289   2.764 0.005708 **
listener_groupmonolingualminus             0.81595    0.35007   2.331 0.019762 *
listener_groupmultilingualminus            0.43941    0.29685   1.480 0.138817
keyword_categoryfd:speaker_nationalityEN  -0.42837    0.14376  -2.980 0.002885 **
keyword_categorytheta:speaker_nationalityEN -0.73707  0.20681  -3.564 0.000365 ***
keyword_categoryfd:ordero2                 0.17686    0.07179   2.464 0.013751 *
keyword_categorytheta:ordero2              0.30911    0.07037   4.393 1.12e-05 ***
```

Monolingual minus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.42569    0.29494   1.443  0.14893
keyword_categoryschwa                     -0.10049    0.13226  -0.760  0.44737
keyword_categorytheta                     -0.01846    0.15214  -0.121  0.90341
speaker_nationalityEN                      0.01541    0.22536   0.068  0.94550
ordero2                                   -0.38328    0.09645  -3.974 7.07e-05 ***
listener_groupmonoandmultiplus            -0.00647    0.27465  -0.024  0.98120
listener_groupmultilingualminus           -0.37672    0.30233  -1.246  0.21274
listener_groupnon_native                  -0.81599    0.35025  -2.330  0.01982 *
keyword_categoryschwa:speaker_nationalityEN  0.42844  0.14374   2.981  0.00288 **
keyword_categorytheta:speaker_nationalityEN -0.30872  0.17681  -1.746  0.08080 .
keyword_categoryschwa:ordero2             -0.17680    0.07179  -2.463  0.01379 *
keyword_categorytheta:ordero2              0.13232    0.07494   1.766  0.07745 .
```

Monolingual minus and th on intercept

(model failed to converge)

Monolingual minus and schwa on intercept

(model failed to converge)

English speakers on intercept: monolingual and multilingual plus and theta

(model failed to converge)

English speakers on intercept: monolingual and multilingual plus and theta

(model failed to converge)

English speakers on intercept: monolingual and multilingual plus and fd

```
Fixed effects:
                                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.434652   0.256432   1.695  0.09008 .
keyword_categoryschwa                      0.328006   0.133185   2.463  0.01379 *
keyword_categorytheta                     -0.327213   0.152484  -2.146  0.03188 *
speaker_nationalityDU                     -0.015580   0.225436  -0.069  0.94490
ordero2                                   -0.383224   0.096450  -3.973 7.09e-05 ***
listener_groupmonolingualminus             0.006299   0.274997   0.023  0.98172
listener_groupmultilingual                -0.370212   0.264637  -1.399  0.16183
listener_groupnon_native                  -0.809418   0.293058  -2.762  0.00575 **
keyword_categoryschwa:speaker_nationalityDU -0.428514   0.143733  -2.981  0.00287 **
keyword_categorytheta:speaker_nationalityDU  0.308801   0.176873   1.746  0.08083 .
keyword_categoryschwa:ordero2             -0.176851   0.071794  -2.463  0.01377 *
keyword_categorytheta:ordero2              0.132271   0.074947   1.765  0.07759 .
```

7.2.10 Re-leveling of the Order 1 Model

Monolingual and multilingual plus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.41070    0.26993   1.522 0.128129
keyword_categoryschwa                     -0.14504    0.14433  -1.005 0.314921
keyword_categorytheta                     -0.15703    0.15128  -1.038 0.299275
speaker_nationalityEN                      0.07701    0.21766   0.354 0.723473
listener_groupmonolingualminus             0.32334    0.31747   1.018 0.308456
listener_groupmultilingualminus           -0.46370    0.29773  -1.557 0.119362
listener_groupnon_native                  -1.07333    0.31131  -3.448 0.000565 ***
keyword_categoryschwa:speaker_nationalityEN  0.45926    0.09988   4.598 4.26e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.03440    0.10492  -0.328 0.742968
```

Monolingual and multilingual plus and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.25371    0.26713   0.950 0.342223
keyword_categoryfd                         0.15704    0.15126   1.038 0.299182
keyword_categoryschwa                      0.01197    0.13932   0.086 0.931509
speaker_nationalityEN                      0.04256    0.21563   0.197 0.843540
listener_groupmonolingualminus             0.32326    0.31737   1.019 0.308401
listener_groupmultilingualminus           -0.46372    0.29763  -1.558 0.119224
listener_groupnon_native                  -1.07337    0.31129  -3.448 0.000564 ***
keyword_categoryfd:speaker_nationalityEN    0.03441    0.10492   0.328 0.742911
keyword_categoryschwa:speaker_nationalityEN  0.49369    0.09615   5.135 2.83e-07 ***
```

Monolingual and multilingual plus and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.26563    0.26307   1.010  0.31262
keyword_categoryfd                         0.14495    0.14433   1.004  0.31522
keyword_categorytheta                     -0.01205    0.13933  -0.087  0.93107
speaker_nationalityEN                      0.53630    0.21340   2.513  0.01197 *
listener_groupmonolingualminus             0.32338    0.31738   1.019  0.30825
listener_groupmultilingualminus           -0.46366    0.29759  -1.558  0.11923
listener_groupnon_native                  -1.07330    0.31107  -3.450  0.00056 ***
keyword_categoryfd:speaker_nationalityEN    -0.45924    0.09988  -4.598 4.27e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.49366    0.09615  -5.134 2.83e-07 ***
```

Multilingual minus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.05298    0.28435  -0.186   0.8522
keyword_categoryschwa                     -0.14502    0.14433  -1.005   0.3150
keyword_categorytheta                     -0.15701    0.15128  -1.038   0.2993
speaker_nationalityEN                      0.07699    0.21767   0.354   0.7236
listener_groupmonoandmultiplus             0.46367    0.29775   1.557   0.1194
listener_groupmonolingualminus             0.78700    0.32986   2.386   0.0170 *
listener_groupnon_native                  -0.60966    0.32388  -1.882   0.0598 .
keyword_categoryschwa:speaker_nationalityEN  0.45925    0.09989   4.598 4.27e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.03441    0.10492  -0.328   0.7429
```

Multilingual minus and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.25371    0.26713   0.950 0.342223
keyword_categoryfd                         0.15704    0.15126   1.038 0.299182
keyword_categoryschwa                      0.01197    0.13932   0.086 0.931509
speaker_nationalityEN                      0.04256    0.21563   0.197 0.843540
listener_groupmonolingualminus             0.32326    0.31737   1.019 0.308401
listener_groupmultilingualminus           -0.46372    0.29763  -1.558 0.119224
listener_groupnon_native                  -1.07337    0.31129  -3.448 0.000564 ***
keyword_categoryfd:speaker_nationalityEN   0.03441    0.10492   0.328 0.742911
keyword_categoryschwa:speaker_nationalityEN  0.49369    0.09615   5.135 2.83e-07 ***
```

Multilingual minus and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.19804    0.27773  -0.713   0.4758
keyword_categoryfd                         0.14502    0.14432   1.005   0.3149
keyword_categorytheta                     -0.01198    0.13932  -0.086   0.9315
speaker_nationalityEN                      0.53627    0.21335   2.514   0.0120 *
listener_groupmonoandmultiplus             0.46366    0.29750   1.559   0.1191
listener_groupmonolingualminus             0.78702    0.32951   2.388   0.0169 *
listener_groupnon_native                  -0.60967    0.32365  -1.884   0.0596 .
keyword_categoryfd:speaker_nationalityEN  -0.45926    0.09988  -4.598 4.27e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.49366    0.09615  -5.134 2.83e-07 ***
```

Non-native and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.66262    0.29865  -2.219  0.02651 *
keyword_categoryschwa                     -0.14503    0.14434  -1.005  0.31499
keyword_categorytheta                     -0.15702    0.15129  -1.038  0.29935
speaker_nationalityEN                      0.07699    0.21769   0.354  0.72360
listener_groupmonoandmultiplus             1.07332    0.31150   3.446  0.00057 ***
listener_groupmonolingualminus             1.39664    0.34228   4.080 4.50e-05 ***
listener_groupmultilingualminus            0.60963    0.32401   1.882  0.05990 .
keyword_categoryschwa:speaker_nationalityEN  0.45926    0.09989   4.598 4.27e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.03440    0.10493  -0.328  0.74302
```

Non-native and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.81965    0.29621  -2.767 0.005655 **
keyword_categoryfd                         0.15704    0.15129   1.038 0.299261
keyword_categoryschwa                      0.01200    0.13934   0.086 0.931370
speaker_nationalityEN                      0.04259    0.21564   0.197 0.843447
listener_groupmonoandmultiplus             1.07330    0.31133   3.447 0.000566 ***
listener_groupmonolingualminus             1.39663    0.34215   4.082 4.47e-05 ***
listener_groupmultilingualminus            0.60964    0.32393   1.882 0.059838 .
keyword_categoryfd:speaker_nationalityEN   0.03440    0.10492   0.328 0.743028
keyword_categoryschwa:speaker_nationalityEN  0.49366    0.09615   5.134 2.83e-07 ***
```

Non-native and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.80767    0.29265  -2.760 0.005783 **
keyword_categoryfd                         0.14504    0.14434   1.005 0.314977
keyword_categorytheta                     -0.01199    0.13935  -0.086 0.931413
speaker_nationalityEN                      0.53625    0.21342   2.513 0.011985 *
listener_groupmonoandmultiplus             1.07332    0.31154   3.445 0.000571 ***
listener_groupmonolingualminus             1.39666    0.34229   4.080 4.50e-05 ***
listener_groupmultilingualminus            0.60965    0.32399   1.882 0.059874 .
keyword_categoryfd:speaker_nationalityEN  -0.45927    0.09989  -4.598 4.27e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.49366    0.09616  -5.134 2.84e-07 ***
```

Monolingual minus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.73400    0.30476   2.408    0.016 *
keyword_categoryschwa                     -0.14506    0.14432  -1.005    0.315
keyword_categorytheta                     -0.15703    0.15127  -1.038    0.299
speaker_nationalityEN                      0.07695    0.21765   0.354    0.724
listener_groupmonoandmultiplus            -0.32322    0.31729  -1.019    0.308
listener_groupmultilingualminus           -0.78691    0.32958  -2.388    0.017 *
listener_groupnon_native                  -1.39659    0.34184  -4.085 4.40e-05 ***
keyword_categoryschwa:speaker_nationalityEN  0.45929    0.09988   4.598 4.26e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.03437    0.10492  -0.328    0.743
```

Monolingual minus and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.57707    0.30232   1.909   0.0563 .
keyword_categoryfd                         0.15698    0.15127   1.038   0.2994
keyword_categoryschwa                      0.01197    0.13933   0.086   0.9315
speaker_nationalityEN                      0.04251    0.21561   0.197   0.8437
listener_groupmonoandmultiplus            -0.32322    0.31731  -1.019   0.3084
listener_groupmultilingualminus           -0.78697    0.32966  -2.387   0.0170 *
listener_groupnon_native                  -1.39669    0.34204  -4.083 4.44e-05 ***
keyword_categoryfd:speaker_nationalityEN   0.03447    0.10492   0.329   0.7425
keyword_categoryschwa:speaker_nationalityEN  0.49370    0.09614   5.135 2.82e-07 ***
```

Monolingual minus and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.58899    0.29909   1.969   0.0489 *
keyword_categoryfd                         0.14504    0.14433   1.005   0.3149
keyword_categorytheta                     -0.01199    0.13935  -0.086   0.9314
speaker_nationalityEN                      0.53623    0.21343   2.512   0.0120 *
listener_groupmonoandmultiplus            -0.32331    0.31763  -1.018   0.3087
listener_groupmultilingualminus           -0.78699    0.32986  -2.386   0.0170 *
listener_groupnon_native                  -1.39663    0.34221  -4.081 4.48e-05 ***
keyword_categoryfd:speaker_nationalityEN  -0.45926    0.09989  -4.598 4.27e-06 ***
keyword_categorytheta:speaker_nationalityEN -0.49367    0.09615  -5.134 2.83e-07 ***
```

7.2.11 Re-leveling of the Order 2 Model

Monolingual and multilingual plus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.11589    0.25285  -0.458  0.64670
keyword_categoryschwa                     -0.33723    0.10326  -3.266  0.00109 **
keyword_categorytheta                      0.18871    0.10784   1.750  0.08012 .
speaker_nationalityEN                      0.23048    0.19593   1.176  0.23946
listener_groupmonolingualminus             0.36936    0.32291   1.144  0.25269
listener_groupmultilingualminus           -0.49837    0.30318  -1.644  0.10021
listener_groupnon_native                  -0.79491    0.31690  -2.508  0.01213 *
keyword_categoryschwa:speaker_nationalityEN  0.49520    0.09825   5.040 4.66e-07 ***
keyword_categorytheta:speaker_nationalityEN -0.42899    0.10030  -4.277 1.89e-05 ***
```

Monolingual and multilingual plus and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.07285    0.25270   0.288   0.7731
keyword_categoryfd                        -0.18872    0.10785  -1.750   0.0801 .
keyword_categoryschwa                     -0.52595    0.10292  -5.110 3.22e-07 ***
speaker_nationalityEN                     -0.19848    0.19609  -1.012   0.3114
listener_groupmonolingualminus             0.36939    0.32319   1.143   0.2531
listener_groupmultilingualminus           -0.49841    0.30347  -1.642   0.1005
listener_groupnon_native                  -0.79500    0.31710  -2.507   0.0122 *
keyword_categoryfd:speaker_nationalityEN   0.42895    0.10031   4.276 1.90e-05 ***
keyword_categoryschwa:speaker_nationalityEN  0.92419    0.09900   9.335  < 2e-16 ***
```

Monolingual and multilingual plus and schwa on intercept

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.45313	0.25074	-1.807	0.070734	.
keyword_categoryfd	0.33724	0.10326	3.266	0.001091	**
keyword_categorytheta	0.52595	0.10292	5.110	3.22e-07	***
speaker_nationalityEN	0.72571	0.19508	3.720	0.000199	***
listener_groupmonolingualminus	0.36937	0.32308	1.143	0.252920	
listener_groupmultilingualminus	-0.49842	0.30343	-1.643	0.100462	
listener_groupnon_native	-0.79494	0.31710	-2.507	0.012179	*
keyword_categoryfd:speaker_nationalityEN	-0.49524	0.09826	-5.040	4.65e-07	***
keyword_categorytheta:speaker_nationalityEN	-0.92420	0.09899	-9.336	< 2e-16	***

Multilingual minus and fd on intercept

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.61431	0.26922	-2.282	0.02250	*
keyword_categoryschwa	-0.33724	0.10326	-3.266	0.00109	**
keyword_categorytheta	0.18871	0.10784	1.750	0.08012	.
speaker_nationalityEN	0.23049	0.19593	1.176	0.23943	
listener_groupmonoandmultiplus	0.49843	0.30340	1.643	0.10042	
listener_groupmonolingualminus	0.86781	0.33600	2.583	0.00980	**
listener_groupnon_native	-0.29653	0.33012	-0.898	0.36905	
keyword_categoryschwa:speaker_nationalityEN	0.49523	0.09825	5.040	4.65e-07	***
keyword_categorytheta:speaker_nationalityEN	-0.42897	0.10030	-4.277	1.90e-05	***

Multilingual minus and th on intercept

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.42559	0.26884	-1.583	0.11341	
keyword_categoryfd	-0.18871	0.10784	-1.750	0.08012	.
keyword_categoryschwa	-0.52596	0.10291	-5.111	3.21e-07	***
speaker_nationalityEN	-0.19849	0.19604	-1.013	0.31130	
listener_groupmonoandmultiplus	0.49844	0.30341	1.643	0.10042	
listener_groupmonolingualminus	0.86779	0.33595	2.583	0.00979	**
listener_groupnon_native	-0.29653	0.33007	-0.898	0.36897	
keyword_categoryfd:speaker_nationalityEN	0.42895	0.10030	4.277	1.90e-05	***
keyword_categoryschwa:speaker_nationalityEN	0.92419	0.09899	9.336	< 2e-16	***

Multilingual minus and schwa on intercept

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.95152	0.26725	-3.560	0.000370	***
keyword_categoryfd	0.33724	0.10326	3.266	0.001092	**
keyword_categorytheta	0.52594	0.10292	5.110	3.22e-07	***
speaker_nationalityEN	0.72569	0.19507	3.720	0.000199	***
listener_groupmonoandmultiplus	0.49844	0.30347	1.642	0.100490	
listener_groupmonolingualminus	0.86776	0.33608	2.582	0.009823	**
listener_groupnon_native	-0.29655	0.33019	-0.898	0.369122	
keyword_categoryfd:speaker_nationalityEN	-0.49524	0.09826	-5.040	4.65e-07	***
keyword_categorytheta:speaker_nationalityEN	-0.92420	0.09899	-9.336	< 2e-16	***

Non-native and fd on intercept

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.91081	0.28475	-3.199	0.001381	**
keyword_categoryschwa	-0.33724	0.10327	-3.266	0.001092	**
keyword_categorytheta	0.18870	0.10784	1.750	0.080165	.
speaker_nationalityEN	0.23046	0.19598	1.176	0.239609	
listener_groupmonoandmultiplus	0.79495	0.31727	2.506	0.012224	*
listener_groupmonolingualminus	1.16428	0.34849	3.341	0.000835	***
listener_groupmultilingualminus	0.29652	0.33029	0.898	0.369327	
keyword_categoryschwa:speaker_nationalityEN	0.49524	0.09826	5.040	4.65e-07	***
keyword_categorytheta:speaker_nationalityEN	-0.42895	0.10031	-4.276	1.90e-05	***

Non-native and th on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.72212    0.28407  -2.542 0.011022 *
keyword_categoryfd                        -0.18870    0.10784  -1.750 0.080143 .
keyword_categoryschwa                     -0.52594    0.10291  -5.110 3.21e-07 ***
speaker_nationalityEN                     -0.19850    0.19604  -1.013 0.311289
listener_groupmonoandmultiplus             0.79495    0.31699   2.508 0.012149 *
listener_groupmonolingualminus             1.16432    0.34831   3.343 0.000829 ***
listener_groupmultilingualminus            0.29654    0.32995   0.899 0.368787
keyword_categoryfd:speaker_nationalityEN   0.42895    0.10030   4.277 1.90e-05 ***
keyword_categoryschwa:speaker_nationalityEN  0.92419    0.09899   9.336  < 2e-16 ***
```

Non-native and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -1.24807    0.28262  -4.416 1.00e-05 ***
keyword_categoryfd                         0.33724    0.10326   3.266 0.001091 **
keyword_categorytheta                      0.52594    0.10292   5.110 3.21e-07 ***
speaker_nationalityEN                      0.72569    0.19507   3.720 0.000199 ***
listener_groupmonoandmultiplus             0.79497    0.31703   2.508 0.012158 *
listener_groupmonolingualminus             1.16433    0.34822   3.344 0.000827 ***
listener_groupmultilingualminus            0.29654    0.33010   0.898 0.369001
keyword_categoryfd:speaker_nationalityEN  -0.49523    0.09825  -5.040 4.65e-07 ***
keyword_categorytheta:speaker_nationalityEN -0.92418    0.09899  -9.336  < 2e-16 ***
```

Monolingual minus and fd on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.25348    0.29121   0.870 0.384068
keyword_categoryschwa                     -0.33724    0.10326  -3.266 0.001092 **
keyword_categorytheta                      0.18871    0.10784   1.750 0.080140 .
speaker_nationalityEN                      0.23047    0.19594   1.176 0.239509
listener_groupmonoandmultiplus            -0.36933    0.32312  -1.143 0.253031
listener_groupmultilingualminus           -0.86775    0.33594  -2.583 0.009793 **
listener_groupnon_native                  -1.16434    0.34823  -3.344 0.000827 ***
keyword_categoryschwa:speaker_nationalityEN  0.49523    0.09825   5.040 4.65e-07 ***
keyword_categorytheta:speaker_nationalityEN -0.42895    0.10030  -4.277 1.90e-05 ***
```

Monolingual minus and theta on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.44219    0.29088   1.520 0.128467
keyword_categoryfd                        -0.18872    0.10783  -1.750 0.080101 .
keyword_categoryschwa                     -0.52597    0.10291  -5.111 3.21e-07 ***
speaker_nationalityEN                     -0.19847    0.19603  -1.012 0.311334
listener_groupmonoandmultiplus            -0.36937    0.32307  -1.143 0.252916
listener_groupmultilingualminus           -0.86781    0.33593  -2.583 0.009786 **
listener_groupnon_native                  -1.16432    0.34828  -3.343 0.000828 ***
keyword_categoryfd:speaker_nationalityEN   0.42895    0.10030   4.277 1.90e-05 ***
keyword_categoryschwa:speaker_nationalityEN  0.92421    0.09899   9.336  < 2e-16 ***
```

Monolingual minus and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.08374    0.28892  -0.290 0.771934
keyword_categoryfd                         0.33725    0.10326   3.266 0.001091 **
keyword_categorytheta                      0.52596    0.10291   5.111 3.21e-07 ***
speaker_nationalityEN                      0.72571    0.19504   3.721 0.000199 ***
listener_groupmonoandmultiplus            -0.36938    0.32284  -1.144 0.252559
listener_groupmultilingualminus           -0.86779    0.33570  -2.585 0.009737 **
listener_groupnon_native                  -1.16431    0.34809  -3.345 0.000823 ***
keyword_categoryfd:speaker_nationalityEN  -0.49524    0.09825  -5.040 4.64e-07 ***
keyword_categorytheta:speaker_nationalityEN -0.92420    0.09899  -9.337  < 2e-16 ***
```

American speakers, monolingual and multilingual plus and schwa on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.27259    0.25092   1.086 0.277316
keyword_categoryfd                        -0.15799    0.10401  -1.519 0.128783
keyword_categorytheta                     -0.39824    0.10141  -3.927 8.59e-05 ***
speaker_nationalityDU                     -0.72569    0.19506  -3.720 0.000199 ***
listener_groupmonolingualminus             0.36935    0.32307   1.143 0.252935
listener_groupmultilingual                -0.49843    0.30334  -1.643 0.100346
listener_groupnon_native                  -0.79497    0.31703  -2.508 0.012156 *
keyword_categoryfd:speaker_nationalityDU   0.49523    0.09825   5.040 4.64e-07 ***
keyword_categorytheta:speaker_nationalityDU  0.92419    0.09898   9.337  < 2e-16 ***
```

American speakers, monolingual and multilingual plus and final devoicing on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                                0.11459    0.25311   0.453   0.6507
keyword_categoryschwa                      0.15799    0.10402   1.519   0.1288
keyword_categorytheta                     -0.24025    0.10695  -2.246   0.0247 *
speaker_nationalityDU                     -0.23046    0.19592  -1.176   0.2395
listener_groupmonolingualminus             0.36934    0.32299   1.144   0.2528
listener_groupmultilingual                -0.49842    0.30328  -1.643   0.1003
listener_groupnon_native                  -0.79499    0.31688  -2.509   0.0121 *
keyword_categoryschwa:speaker_nationalityDU -0.49522    0.09825  -5.040 4.65e-07 ***
keyword_categorytheta:speaker_nationalityDU  0.42896    0.10030   4.277 1.89e-05 ***
```

American speakers, monolingual and multilingual plus and theta on intercept

```
Fixed effects:
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.12563    0.25212  -0.498   0.6183
keyword_categoryfd                         0.24025    0.10695   2.246   0.0247 *
keyword_categoryschwa                      0.39824    0.10141   3.927  8.6e-05 ***
speaker_nationalityDU                      0.19850    0.19608   1.012   0.3114
listener_groupmonolingualminus             0.36936    0.32316   1.143   0.2531
listener_groupmultilingual                -0.49845    0.30353  -1.642   0.1006
listener_groupnon_native                  -0.79497    0.31713  -2.507   0.0122 *
keyword_categoryfd:speaker_nationalityDU  -0.42896    0.10030  -4.277  1.9e-05 ***
keyword_categoryschwa:speaker_nationalityDU -0.92419    0.09899  -9.336  < 2e-16 ***
```