L2 PHONEME CATEGORIES, LEXICAL ACCESS AND FOREIGN ACCENT

Frank Lorenz

Universität Erfurt, Germany
frank.lorenz@uni-erfurt.de

ABSTRACT

An experiment using a visual world eye-tracking paradigm seeks to shed light on the relationship between native-like vs. foreign-accented speech and the phoneme categories of the speakers.

Previous research has shown that competitor images are fixated longer a) with a similar phonetic onset in the speaker's L1 and b) with an L2 vowel that is not a phoneme category in the speaker's L1. This study presented target images containing [æ] in English to two groups of native German speakers: with and without a distinct German accent in English.

English /æ/ and /ɛ/ may be conflated to /ɛ/ by German speakers. It was investigated whether speakers without foreign-accented speech showed fewer fixations to competitors containing [ɛ] in German, as they acquired a new phoneme category. Secondly, it was examined whether the two groups differ with competitor images containing [a] in German and how this could be interpreted.

Keywords: eye tracking, phoneme categories, learner phonology, foreign accent

1. INTRODUCTION

Phoneme categories are developed early in life. While infants are able to discriminate any phonetic contrast in the first six months of their life, they lose this ability in the subsequent six months in favour of contrasts that are phonemic in their L1. [16] This has an impact on L2 acquisition where contrasts that are not phonemic in the L1 have to be perceived and produced. Current models of speech perception and production assume a strong link between L1 phonology and its impact on L2 acquisition. The Speech Learning Model (SLM) proposes that L1 and L2 phonetic subsystems share a common phonological space. If an L2 sound is sufficiently distinct from the closest L1 phoneme, a new category will be formed for this sound. If it is perceived as similar to an L1 sound, however, there will be a merged L1/L2 category. [12] The extended Perceptual Assimilation Model (PAM-L2) claims that listeners will perceptually assimilate an L2 sound to the closest L1 phoneme category. It is possible that two L2 phonemes are assimilated to the same L1 category, either as equally good exemplars or with a difference in goodness of fit. If discrimination of the assimilated L2 phones is possible, it is likely that for the worse fit a new phonetic and phonological category is developed. [2]

Different learners may succeed differently well in the task of establishing L2 categories. The present study uses an eye tracking experiment to examine the link between a foreign or native-like accent and the speakers' potential phoneme categories. Eye tracking studies over the past decade have helped us understand how speakers access lexical items, in what way phoneme categories may be organised and in the case of bilingual speakers which languages are accessed in the perception process. Previous research with the help of visual world designs has revealed two essential findings.

For bilingual speakers, the speaker's L1 is also activated in the comprehension process [13, 18]. Weber and Cutler [18], for instance, have shown that in an experiment conducted entirely in English, Dutch participants showed increased fixations to objects when the lexical items shared the same phonetic onset in their L1 Dutch with the English target word. (With the auditory stimulus *kitten*, a *box* that translates as *kist* in Dutch is fixated more often and until later.)

Secondly, competitors receive more fixations when the target contains a sound that is not a phoneme in the speaker's L1 and is hence confusable with the sound of the competitor [6, 9, 18]. (When told to click the *panda*, participants showed increased fixations to a *pencil*, as the English /æ/-/ɛ/ contrast is confusable to Dutch speakers.)

The present study amalgamates these two designs by presenting German native speakers with English target words that contain [æ] and competitor items that are phonetically dissimilar in English, but have a potentially confusing vowel in their German translation.

The English /æ/-/ɛ/ phoneme contrast is not established in German and speakers typically pronounce /æ/ as [ɛ]. (Unlike other languages, such as Spanish, /æ/ is not mapped to /a/.) It could be hypothesised that German speakers learning English initially have one phoneme category that comprises both [æ] and [ɛ], but as their language acquisition progresses, they develop a separate phoneme category for /æ/. If we assume that speakers without

a foreign accent have mastered the task of acquiring L2 phoneme categories better than speakers with a distinct foreign (L1) accent, then the former should show fewer fixations when presented with a competitor containing [ɛ] in the German translation than the latter.

German learners who work to improve their perception and production of the /æ/-/ɛ/ contrast sometimes comment that to them [æ] 'somehow sounds like' [a]. This is not unfounded, since German /a/ is the closest L1 phoneme category and it is also supported by spelling, since both English [æ] and German [a] are always spelled ⟨a⟩ and although spoken language has primacy over its written form, it has been shown that spelling has an effect in auditory word recognition tasks, for instance. [19] This, however, raises the question whether advanced learners did in fact acquire a new phoneme category for /æ/ or whether they re-assimilated [æ] as an allophone of German /a/. If that were the case, it would be of interest whether competitors containing [a] received more fixations than distractors that contain a different vowel.

2. METHOD

2.1. Participants

Participants were German native speakers and students of the BA programme English and American Studies at the University of Erfurt in their first year of study. All students of a compulsory phonetics lecture were asked to record a list of five sample sentences at home that contained among others [æ] and [ɛ] with whatever means they had available. They were told that the purpose of the recording was to evaluate students' pronunciation in an anonymised way, but that they might be invited to an experiment later.

The sample recordings were played to two English native speakers (one British and one American). They were asked to rate the accent of the speakers on a five-point Likert scale from "very strong German accent" to "no foreign accent". Raters were instructed that many students would most likely fall in the middle category and that this category should represent what to them is a typically average pronunciation of a German language learner.

Students who were given a higher or lower than average rating by both judges on their foreign-accentedness were selected for the study. In all, 44 students took part in the experiment, 22 with a distinct German accent and 22 without or with a distinctly mild foreign accent. It would have been desirable to have more participants in each group, but as this would have meant inviting students with an average rating on their accent, it would have skewed the results. As much more female students took the course and submitted a recording, only female students were invited to the experiment to avoid any gender bias.

2.2. Materials

Participants were presented with four images and were asked to look at one of the four items. The target items were 24 words that contained [æ] in English, but whose German translation was phonetically dissimilar in that it did not contain the same onset plus short [a] (e.g. *cat*/*Katze* was rejected).

These targets were paired with two kinds of competitors. On the one hand, with items that had the same phonetic onset and contained [ɛ] in their German translation, but would be phonetically dissimilar in English (e.g. *ham/Schinken–Hemd/shirt*). On the other hand, with items that contained [a] in their German translation, but would again be phonetically dissimilar in English (e.g. *ham/Schinken–Handtuch/towel*). Ideally, also the consonants following the vowels were identical or phonetically similar in target and competitor, but this could not always be guaranteed. For 7 out of the 24 targets, no [ɛ]-competitor could be found. For one target, there was no [a]-competitor. Where targets had both kinds of competitors, participants only saw one competitor, while a different subgroup was presented with the other.

Most images were photos taken from the [3] picture set. Several images were added where no photo was available. All photos were rendered black-and-white to avoid attention to colour. The images had a size of about 5.5 × 5.5 cm. Targets and competitors were put in different positions in all trials, so there would be no learning effect or bias for the top left position. In addition to the 24 test trials that every participant was presented with, 25 distractor trials were created where the target item contained vowels different than [æ].

2.3. Procedure

Students were invited to the eye tracking laboratory. The blinds can be closed in that room, so that there is only one light source. The eye tracker by SMI is mounted below the screen. All students were greeted in English and they were given time to make themselves comfortable. It was explained to them that they would hear a word over headphones and should look at the corresponding picture. It was pointed out that they might not know the word they hear and should then indicate this. These trials were later discarded.

After a calibration phase at the beginning of the experiment, a cross was displayed in the middle of the

screen for 4000 ms which students had been instructed to fixate. Afterwards, the cross disappeared and the four images (target, competitor and two distractors) were displayed on the screen at equal distances from the cross. Simultaneously, a recording of the target item was played over headphones that had been recorded by a speaker with a General British accent. After 4000 ms, the images disappeared and the cross reappeared to start the next trial. The different trials were displayed in randomised order.

After the experiment, the instructor switched to German. All participants were shown pictures of the competitor items again and they were asked to name these items in German. Where participants provided a different word than anticipated, this was noted and these trials were later discarded.

3. RESULTS

For the analysis, only fixations to the four areas of interest were analysed. Fixation proportions were calculated starting at 200 ms after the beginning of the trial in steps of 20 ms until 1000 ms. Fixations to the two distractors were averaged. The figures below show fixation proportions for all participants and all test trials, except the ones that had to be removed because the participant did not know the target word or labelled the competitor differently.

When presented with competitor items containing [ɛ] in their German translation, fixation proportion differ from distractor items with different vowels in both groups.

The group with a distinct German accent shows nearly twice as many fixations to competitors containing [ɛ] than to the target from the onset until 460 ms. Afterwards, as looks to the target increase rapidly after 620 ms, the competitor receives more fixations than the distractors until after 800 ms.

The result for the group without a foreign accent is puzzling. The point of target identification is earlier at about 400 ms. However, competitors with [ɛ] show fewer fixations than distractor items: almost half as many fixations, the opposite relation as for the foreign-accent group. (It has to be admitted that this group had the lowest absolute number of fixations and that this finding might be an artefact.)

Figure 1: All trials with an [ɛ]-competitor by participants with a distinct German accent

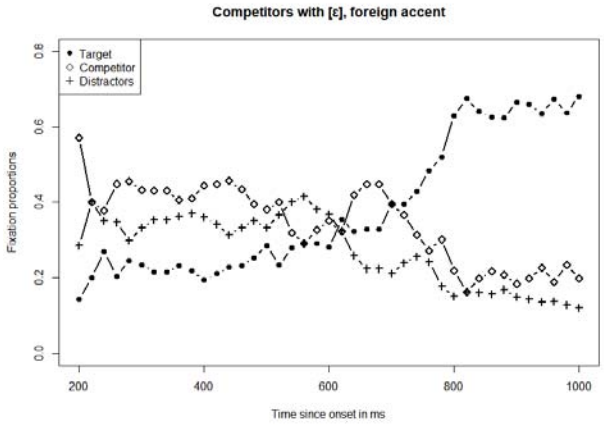
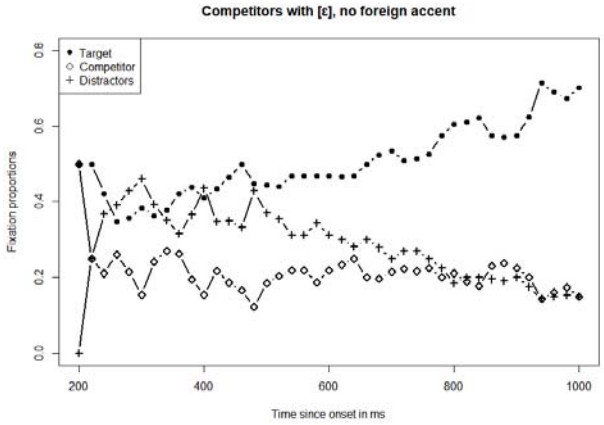


Figure 2: All trials with an [ɛ]-competitor by participants with no foreign accent



The differences between target, competitor and distractor items are not as pronounced, when the competitor items contained [a] in their German translation.

The group with a distinct German accent when presented with competitors containing [a] showed more fixations to the competitor than to the distractor items between 460 ms and 860 ms.

Exploration of the group without a foreign accent shows a similar effect, however not so pronounced. The point of target identification is a little earlier. Furthermore, fixations to the competitor reach the proportion level of distractors about 180 ms earlier. A one-factor ANOVA on the fixation proportions for both groups, however, showed no significant differences between the groups. Further statistical analyses more appropriate to time-course data will have to provide further insights.

Figure 3: All trials with an [a]-competitor by participants with a distinct German accent

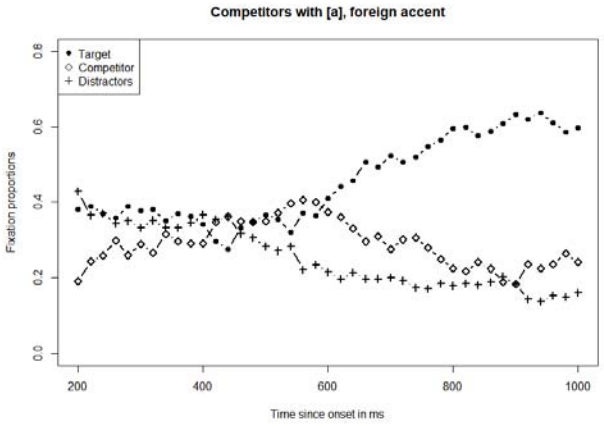
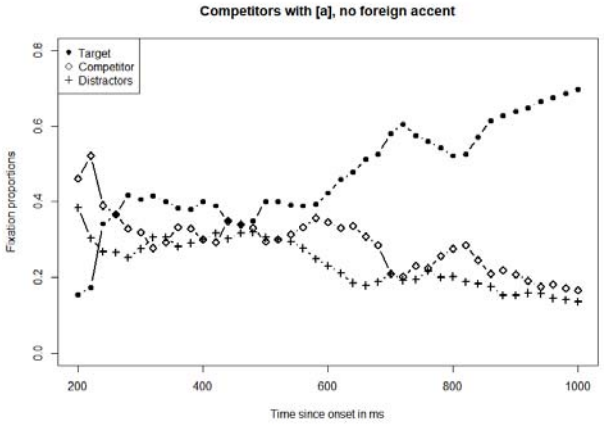


Figure 4: All trials with an [a]-competitor by participants with no foreign accent



4. DISCUSSION

The comparison of two groups of speakers, with and without foreign-accented speech, has revealed some interesting differences. As expected, speakers with a distinct German accent showed more fixations to lexical items when these contained [ɛ] in their German translation. This could be seen as an indication that the hypothesis holds true that speakers with a distinct German accent treat [æ] and [ɛ] as belonging to the same phoneme category, while speakers without a foreign accent have acquired a new phoneme category.

A similar effect was found for competitors that contained [a] in their German translation. These received more fixations than distractor items. The effect was stronger for speakers with foreign-accented speech. An explanation is more difficult here. Words that contain /æ/ in English are not typically pronounced with [a] by German native speakers, although learners working to improve their /ɛ/-/æ/ contrast sometimes comment that [æ] somehow sounds like [a] to them. It could be hypothesised that speakers treat [æ] as an allophone of their L1 /a/-category. However, what speaks against this is that speakers with a foreign accent show higher fixations to competitors both with [ɛ] and [a]. An alternative explanation would be an influence of spelling. For instance, in auditory word recognition experiments, consistent or non-consistent spelling did impact word recognition. [19] It would be conceivable that also in a visual world paradigm orthography does play a role. After all, while not every ⟨a⟩ is pronounced [æ], every [æ] is represented by ⟨a⟩ in spelling and this in turn always corresponds to [a] or [aː] in German.

The findings paint an interesting picture on the organisation and development of phoneme categories of L2 learners. Further research with a higher number of participants is certainly needed to validate these findings. This would also allow for a more fine-grained per item analyses.

5. REFERENCES

[1] Bergmann, C., Nota, A., Sprenger, S., Schmid, M. 2016. L2 immersion causes non-native-like L1 pronunciation in German attriters. *J. Phon.* 58, 71-86. doi:10.1016/j.wocn.2016.07.001.

[2] Best, C., Tyler, M. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In: Bohn, O.-S., Munro, M. J. (eds.), *Language Experience in Second Language Speech Learning*. Amsterdam: John Benjamins. doi:10.1075/lllt.17.07bes.

[3] Brady, T. F., Konkle, T., Alvarez, G. A., Oliva, A. 2008. Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci. USA* 105 (38), 14325-14329.

[4] Cutler, A., Davis, C. 2012. An orthographic effect in phoneme processing, and its limitations. *Front. Psychol.* 3, 18. doi:10.3389/fpsyg.2012.00018.

[5] Cutler, A., Weber, A. 2007. Listening experience and phonetic-to-lexical mapping in L2. *Proc. 16th ICPhS* Saarbrücken, 43-48.

[6] Cutler, A., Weber, A., Otake, T. 2006. Asymmetric mapping from phonetic to lexical representations in second-language listening. *J. Phon.* 34(2), 269-284. doi:10.1016/j.wocn.2005.06.002.

[7] Cutler, A., Weber, A., Smits, R., Cooper, N. 2004. Patterns of English phoneme confusions by native and non-native listeners. *J. Acoust. Soc. Am.* 116(6), 3668-3678. doi:10.1121/1.1810292.

[8] Duyck, W., Van Assche, E., Drieghe, D., Hartsuiker, R. 2007. Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 33. doi:10.1037/0278-7393.33.4.663

[9] Escudero, P., Hayes-Harb, R., Mitterer, H. 2008. Novel second-language words and asymmetric lexical access. *J. Phon.* 36(2), 345-360. doi:10.1016/j.wocn.2007.11.002.

[10] Escudero, P., Simon, E., Mitterer, H. 2012. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *J. Phon.* 40, 280-288. doi:10.1016/j.wocn.2011.11.004

[11] FitzPatrick, I., Indefrey, P. 2010. Lexical competition in nonnative speech comprehension. *J. Cogn. Neurosci.* 22, 1165-1178. doi:10.1162/jocn.2009.21301.

[12] Flege, J. E. 2007. Language contact in bilingualism: Phonetic system interactions. In: Cole, J., Hualde, J. I. (eds.), *Laboratory Phonology* 9. Berlin: Mouton de Gruyter.

[13] Marian, V., Spivey, M. 2003. Competing activation in bilingual language processing: Within- and between-language competition. *Biling. Lang. Cogn.* 6 (2), 97-115. doi:10.1017/S1366728903001068

[14] McQueen, J. M., Cutler, A., Norris, D. 2006. Phonological abstraction in the mental lexicon. *Cogn. Sci.* 30(6), 1113-1126. doi:10.1207/s15516709cog0000_79.

[15] Reinisch, E., Weber, A., Mitterer, H. 2013. Listeners retune phoneme categories across languages. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 75-86. doi:10.1037/a0027979.

[16] Saffran, J. R., Werker, J. F., Werner, L. A. 2006. The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language. In: Damon, W., Lerner, R. M., Kuhn, D., Siegler, R. (eds), *Handbook of Child Psychology*. doi:10.1002/9780470147658.chpsy0202.

[17] Wagner, A. 2008. *Phoneme inventories and patterns of speech sound perception*. PhD Thesis, Radboud University Nijmegen.

[18] Weber, A., Cutler, A. 2004. Lexical competition in non-native spoken-word recognition. *J. Mem. Lang.* 50(1), 1-25. doi:10.1016/S0749-596X(03)00105-0.

[19] Ziegler, J.C., Ferrand, L., Montant, M. 2004. Visual phonology: the effects of orthographic consistency on different auditory word recognition tasks. *Mem. Cognit.* 32(5), 732-41.