

The perception of manipulated German vowels by Turkish GFL learners

Katharina Nimz, Humboldt University of Berlin

The following experiment was to test whether quantity (duration) and/or quality (spectral properties) is the difficult feature for Turkish German as a Foreign Language (GFL) learners to distinguish when learning German vowels. A perception experiment has been conducted in both Turkey and Germany in which Turkish GFL learners (N=22) and German native speakers (N=21) participated in a discrimination task with manipulated nonsense words (Ø age Turkish group: 17.5 (SD=.5); 6 female, 14 male participants. Ø age German group: 17.9 (SD=.7); 12 female, 8 male participants).

Because both durational and spectral features play a role in the discrimination of German long/short vowel pairs (Kohler 1995), a simple discrimination experiment for these pairs would not yield information as to which feature – quality and/or quantity – may be problematic for Turkish GFL learners. For this reason, the German long vowels /i:/, /u:/, /a:/ and the German short vowels /ɪ/, /ʊ/, /a/ were manipulated in reminiscence of a design used by Sendlmeier 1981.

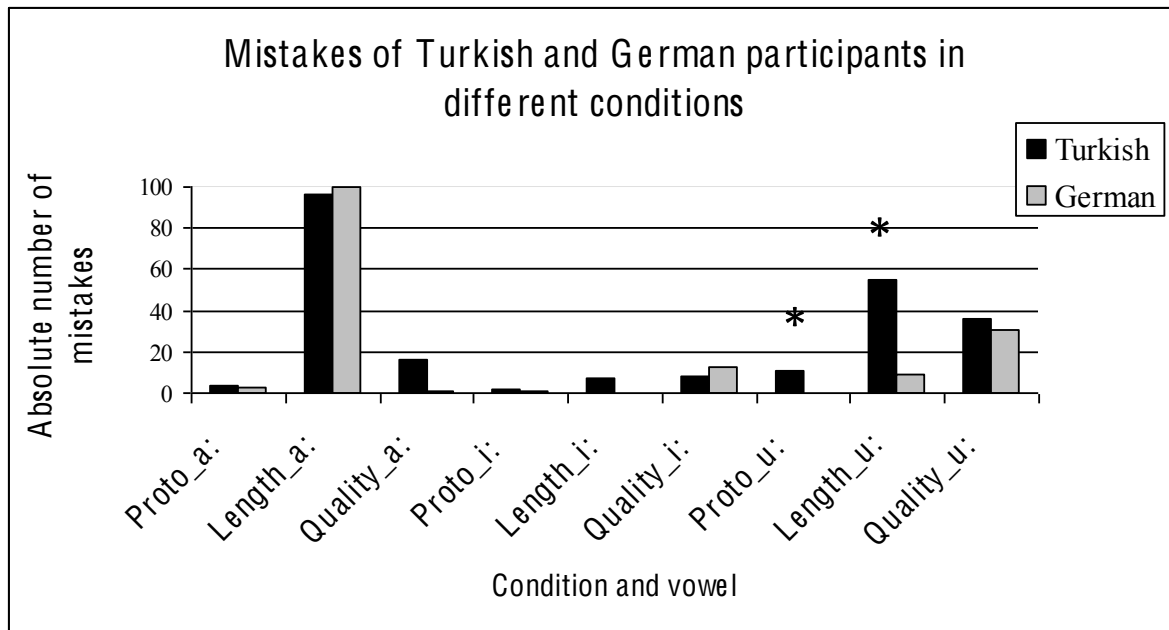
With the help of PRAAT, a prototypical long vowel (spoken by a female German native speaker) was shortened to the average length of its corresponding short counterpart, whereas a prototypical short vowel was lengthened to the average length of its corresponding long counterpart. For the discrimination task, nonsense word pairs were then matched for three conditions: (1) condition “proto”: non-manipulated long vowel vs. non-manipulated short vowel, e.g. /bu:p/ vs. /bop/, (2) condition “length”: lengthened short vowel vs. non-manipulated long vowel, e.g. /bʊ:p/ vs. /bu:p/, (3) condition “quality”: shortened long vowel vs. non-manipulated short vowel, e.g. /bup/ vs. /bʊp/. Each vowel pair was judged 5 times in each condition for being “same” or “different”, plus the control condition “clearly different” (/bap/ versus /bɪp/) and 20 filler pairs (“same”) such as /bu:p/ vs. /bu:p/. In all, each subject rated 80 nonsense word pairs, presented in blocks of 8. The experiment lasted about 5 minutes.

The results (s. next page) indicate that, as expected, almost all instances of the manipulated “a-pair” (= /a-/a:/ vowel pair) in the condition “length” were rated wrongly as “same”, as this pair – in contrast to the others – is distinguished solely on the basis of duration (Weiss and Wängler 1975). Hence, when /a/ is lengthened, it should sound like a true /a:/, which was the case in both the Turkish and the German group (s. Figure 1). However, since group differences were of interest only, the programming language “R” (R Development Core Team 2006) was used to calculate in which condition and for which vowel pair the Turkish participants differed from the German control group. As can be inferred from Figure 1, significant differences could be detected for the “u-pair”: For the condition “length”, the group difference was highly significant with $p < .001$ (U=48), for the condition “proto”, the group effect was significant at $p = .003$ (U=120).

Hence, Turkish participants – in comparison to the German control group – only had obvious difficulties keeping spectral differences apart, as nonsense words in condition “length” had the same length but different quality, i.e. spectral properties. Further, condition “proto” yielded a significant group effect; however, this was only true for the “u-pairs”. Interestingly, none of the “quality” conditions showed significant differences between groups, which suggests that the duration feature does not pose a prominent problem for Turkish learners of L2 German, at least not more so than for German native speakers. This finding supports assumptions made by Bohn 1995 who claims that duration are easy to access whether listeners have had specific linguistic experience with them or not.

Open questions however remain and should be open for discussion: How important is the investigation of *segments* in foreign language acquisition in comparison to *suprasegmental features*? In how far are the results of an experiment with *manipulated vowels* transferable to *everyday speech*? Did the *design of the experiment* help the participants discriminate between vowels that they usually would not be able to distinguish?

Figure 1: Results perception experiment^{1,2}



1. Due to the problem of multiple testing, α was adjusted according to the Bonferroni correction, hence, α was set at .006.
2. It needs to be noted that the group effect in condition “Proto_u” would not be significant if only one German participant had made one mistake in this condition or if only one Turkish participant had made one mistake less. For this reason the effect should be interpreted with caution and I will refrain from its discussion in the current paper.

References:

Bohn, O.-S. 1995: Cross-Language speech perception in adults. First language transfer doesn't tell it all. In: Strange, W. (ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 279-304.

Kohler, K. 1995: *Einführung in die Phonetik des Deutschen*. Berlin: Schmidt.

R Development Core Team. 2006. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>

Sendlmeier, W. F. 1981: Der Einfluß von Qualität und Quantität auf die Perzeption betonter Vokale des Deutschen. *Phonetica* 38, 291-308.

Weiss, Rudolf & Wängler, Hans-Heinrich (1975): Experimental approach to the study of vowel perception in German. *Phonetica* (32), 180-199.