Phonological emergence in Dutch

Phonological emergence in Dutch: Relating perception and production in contact-induced change

Alexander Martin[1], Marieke van Heugten[2], René Kager[3], Sharon Peperkamp[1]

[1]Laboratoire de Sciences Cognitives et Psycholinguistique (ENS, EHESS, CNRS) École Normale Supérieure – PSL Research University

[2]University at Buffalo, The State University of New York

[3]Utrecht Institute of Linguistics OTS Utrecht University

Recent work has established that phonological contrasts are often gradient rather than categorical (e.g., Hall, 2013). One source of this gradience comes from emergent contrasts, such as those created during language change. Here we consider the relation between perception and production in contrasts introduced through borrowing. While there is evidence for perception leading internal sound change (Coetzee, Beddor, & Wissing, 2014; Harrington, Kleber, & Reubold, 2008; Janson, 1983; Pinget, 2015), little is known about the role of perception in *contact*-induced phonological change. In the present study, we explore the loanword adaptation process, taking a contact-induced case of phonological emergence in Dutch as an example.

Over time, Dutch has borrowed many words from neighboring languages, including over 1,300 words from English since 1900 (van der Sijs, 2002). Many of these contain the foreign sound /g/, which the Dutch inventory lacks. Such words may be produced in three ways. First, as the grapheme <g> corresponds to the Dutch fricative /χ/ (and its regional variants), speakers may produce the <g> as a fricative if they follow an orthographic strategy of adaptation. Second, following a perceptually-shaped phonetic strategy, they may produce it using the Dutch sound that most closely resembles it, namely [k]. And third, they may pronounce it as a voiced stop [g], signaling a change in the sound system of the language; we will refer to this as the phonological strategy. This last strategy has had as an effect that /g/ has been slowly making its way into the language, to the extent that now even a minimal pair between native /k/ and emerging /g/ exists: /koːl/, *cabbage* ~ /goːl/, *goal*. Previous research on the loan phoneme /g/ in Dutch has indicated regional variation in its production, with evidence for the use of all three strategies defined above (van Bezooijen & Gerritsen, 1994; van de Velde & van Hout, 2002), but these studies have only considered a handful of loanwords, with a reduced set of speakers.

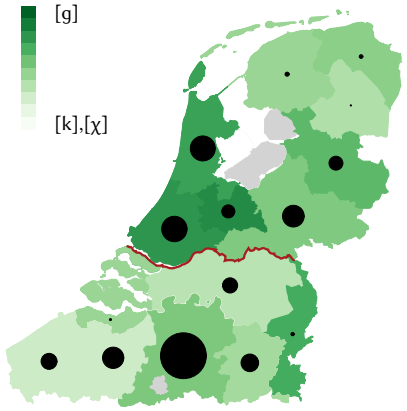

Figure 1: Mean [g] production in loanwords containing /g/ across the Netherlands and Flanders. Black circles represent the number of data points in each region.

The present study examines both the production of loanwords with /g/ and the ability to perceive the emerging contrast (/k/~/g/) by native speakers of Dutch. We began our investigation by exploiting the *Corpus Gesproken Nederlands* (Oostdijk, 2000), a large collection of recordings of spoken Dutch from across the Netherlands and Flanders (the Dutch-speaking part of Belgium). We extracted a total of 634 tokens of 49 loanwords originating from a variety of languages containing /g/ produced by 355 Dutch speakers (211 men) with a variety of regional backgrounds. Extracted tokens were coded as containing an adapted sound (i.e., [χ] or [k]), or the emerging sound [g]. Results are plotted on the map in Figure 1, with darker regions representing more [g] productions. Note that regions in the north (darker) show more [g] production than regions in the south (β=-1.38, SE=0.25, $\chi^2(1)$=34.01, p<0.0001). This split roughly follows a historical division referred to in Dutch as the "great rivers", which we use in the present study to divide the area dialectically between north and south, shown as a red line in Figure 1 (Donaldson, 1983).

Next, we tested 51 participants at Utrecht University (located in the darkest region of Figure 1), 29 of whom spent the majority of their youth in the north (darker regions), and 22 in the south (lighter regions). They performed three tasks, one in production and two in perception, and completed a detailed questionnaire concerning socio-economic status and foreign language experience. The northern and southern groups were matched for practically all recorded indices.

In the production task, participants read 60 sentences out loud, 12 of which were targets containing loanwords with one or more /g/s (e.g., *buggy* or *Google*). All of the other sentences contained frequency-matched loanwords, so as not to draw participants' attention to the /g/ sentences. We calculated the proportion of [g] productions by participant (total number of /g/s=696) and found that, in line with the corpus data, southerners produced more adapted tokens (i.e., with [k] or [χ]) than northerners did (β = -0.68, SE = 0.29, $\chi^2(1)$ = 7.72, p < 0.01). Overall, and unlike in the corpus, there were few [χ] productions (less than 10%, compared to 39% in the corpus), indicating that words were borrowed according to the way they sound rather than the way they are written. The lack of fricative productions may be related to the young age of participants tested (range: 18–39, mean: 22.7; compared to the corpus, range: 15–84, mean: 38.1); previous work on Dutch /g/ has indeed indicated possible generational differences (Hamann & de Jonge, 2015).

The perception tasks assessed participants' discrimination of the emerging /k/~/g/ contrast compared to a native control contrast, /p/~/b/. One task used an implicit discrimination paradigm with a syllable identification task (Navarra, Sebastián-Gallés, & Soto-Faraco, 2005). Participants heard disyllabic non-words, whose initial syllable was always either [fu] or [ni], while the second syllable was one of [ka:ms], [ga:ms], [pa:ms], and [ba:ms]. Multiple tokens produced by a female speaker were used. In consistent blocks, the second syllable of the non-word was always the same; in mixed blocks, it was variable. If participants can perceive the contrast used in the second syllable, response times in mixed blocks should be longer than those in consistent blocks. We found that participants were indeed slower to respond in mixed blocks than in consistent blocks (β = -0.05, SE = 0.01, $\chi^2(1)$ = 19.4, p < 0.01), but no effect of contrast (native vs emerging; $\chi^2(1) < 1$), nor an interaction ($\chi^2(1) < 1$) was observed. This indicates that participants perceived the /k/~/g/ and /p/~/b/ contrasts equally well. Moreover, no effect of region of origin was found ($\chi^2(1) < 1$), indicating that Northerners and Southerners alike could perceive both the native and emerging contrasts.

The second perception task used an explicit ABX discrimination paradigm, in which participants heard sequences of two minimally different Dutch pseudowords, followed by a third one that was identical to either the first or the second (e.g. ['mœŷ**k**ərəχ] - ['mœŷ**g**ərəχ] - ['mœŷ**k**ərəχ]). The A and B tokens were produced by two different female voices, and the X token by a male voice, promoting phonological rather than simple acoustic processing. Half of the trials tested perception of the emerging /k/~/g/ contrast, and the other half of the native /p/~/b/ contrast. Accuracy was very high overall but marginally higher for the native than for the emerging contrast (native: 91%, emerging: 88%; β = -0.42, SE = 0.21, $\chi^2(1)$ = 3.68, p = 0.055). As in the previous task, no effect of region of origin was found ($\chi^2(1) < 1$).Thus, in the ABX but not the syllable identification task, participants showed a trend towards better performance at the native than the emerging contrast, potentially due to a lack of variability in the stimuli used in the implicit task (only one speaker, compared to three in the ABX). Results from both tasks, though, suggest that both contrasts are perceived.

Finally, in order to investigate the relation between perception and production, we compared each participant's performance across the different tasks. We first calculated a difference score in each perceptual task: For the syllable identification task, this score was defined as the difference between the mixed and consistent blocks for the native contrast minus the same difference for the emerging contrast. For the ABX task, it was defined as the difference in accuracy between the native and emerging contrasts. For both tasks, a positive difference score indicates better performance on the native contrast. We then compared the difference scores to the proportion of /g/ production in the production task. Regressions for the two perceptual tasks compared to the participants' productions can be seen in Figure 2 and Figure 3. We performed a multiple regression for each perceptual task with the predictors Difference Score and Region. For the implicit task, only an effect of Region was observed (t = -2.66, p = 0.01), while for the ABX task, we observed an interaction (t = -2.13, p = 0.04) between the two factors. We then conducted restricted analyses on the ABX data, performing separate regressions by region. We found no correlation between the perception and production in the Northern group (r^2 = 0.02, p = 0.24), but a highly significant correlation in the Southern group (r^2 = 0.34, p < 0.01), suggesting a difference in the establishment of the emerging contrast in the two populations. Indeed recent work on Dutch dialectology has suggested that sound change can progress differentially in the different regions of the Low Countries (Pinget, 2015). It is thus possible that the emergence is rather more complete in the North than in the South, where we can still observe high degrees of variability in perception.

The results from our corpus study and production task show abundant evidence for the representation of emerging /g/ in Dutch speakers. Furthermore, the production of /g/ seems to be modulated to some extent by participants' ability to *perceive* the contrast; Southerners who performed better in ABX discrimination were more likely to use /g/ when producing loanwords. Overall, our results suggest that the emerging sound /g/ has become well anchored in Dutch, with listeners both perceiving and producing it in contrast with native /k/, though this may be modulated by region. We will discuss social and linguistic factors that might contribute to this change, and specifically potential sources of regional variation.

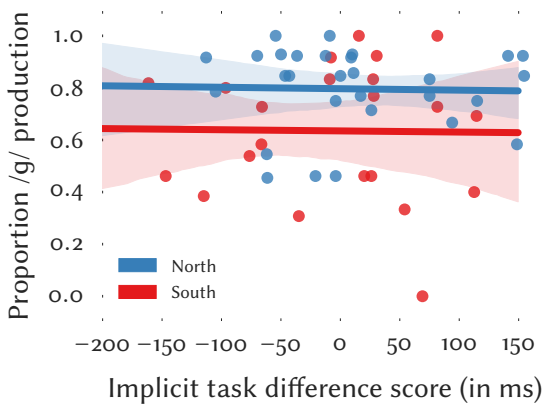

Figure 2: Proportion of /g/ responses by participant as a function of their difference score in the implicit task. Each point represents a participant; those from the North are in blue, the South in red.

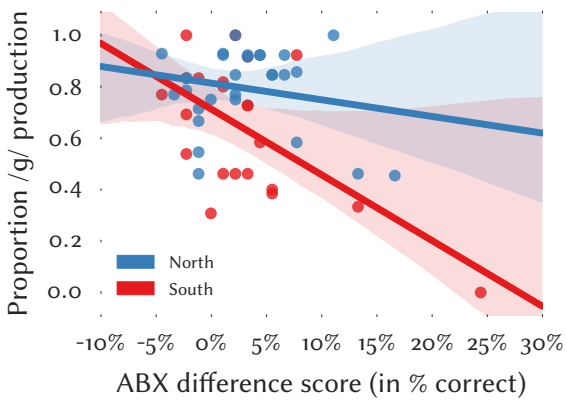

Figure 3: Proportion of /g/ responses by participant as a function of their difference score in the ABX task. Each point represents a participant; those from the North are in blue, the South in red.

References

Coetzee, A. W., Beddor, P. S., & Wissing, D. P. (2014). Emergent tonogenesis in Afrikaans. *The Journal of the Acoustical Society of America*, *135*(4), 2421–2422.

Donaldson, B. C. (1983). *Dutch: a Linguistic History of Holland and Belgium*. Leiden: Uitgeverij Martinus Nijhoff.

Hall, K. C. (2013). A typology of intermediate phonological relationships. *The Linguistic Review*, *30*(2), 215–276.

Hamann, S., & de Jonge, A. (2015). Eliciting the Dutch loan phoneme /g/ with the Menu Task. In *Proceedings on the 18th International Congress of Phonetic Sciences*. Glasgow.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: an acoustic and perceptual study. *The Journal of the Acoustical Society of America*, *123*(5), 2825–2835. doi:10.1121/1.2897042

Janson, T. (1983). Sound change in perception and production. *Language*, *59*(1), 18–34.

Navarra, J., Sebastián-Gallés, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: new evidence from an implicit measure. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(5), 912. doi:10.1037/0096-1523.31.5.912

Oostdijk, N. (2000). The Spoken Dutch Corpus: Overview and first evaluation. *Proceedings of Second International Conference on Language Resources and Evaluation (LREC)*, 887–894.

Pinget, A.-F. (2015). *The actuation of sound change*. Universiteit Utrecht.

van Bezooijen, R., & Gerritsen, M. (1994). De uitspraak van uitheemse woorden in het Standaard-Nederlands: een verkennende studie. *De Nieuwe Taalgids*, *87*(2), 145–161.

van de Velde, H., & van Hout, R. (2002). Uitspraakvariatie in leenwoorden. *NVT-Onderwijs En -Onderzoek in Franstalig Gebied*, *1*, 77–95.

van der Sijs, N. (2002). *Chronologisch woordenboek: de ouderdom en herkomst van onze woorden en betekenissen* (2nd ed.). Amsterdam: LJ Veen.