Comparing Lexically-guided Perceptual Learning in Younger and Older Listeners

Odette Scharenborg[1,2] and Esther Janse[3,1,2]

[1]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[2]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

[3]Centre for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands

Author note

The research by Odette Scharenborg is sponsored by the Max Planck International Research Network on Aging. The research by Esther Janse is supported by a Vidi-grant from the Netherlands Organization for Scientific Research (NWO). We thank the student-assistants of the Comprehension Group and the Adaptive Listening Group of the Max Planck Institute for Psycholinguistics for assistance in preparing and running these experiments, and Marijt Witteman for recording the stimuli. Finally, we thank James McQueen for valuable comments on an earlier version of this manuscript.

Corresponding author:

Odette Scharenborg

Max Planck Institute for Psycholinguistics

Wundtlaan 1

6525 XD Nijmegen

The Netherlands

E-mail: Odette.Scharenborg@mpi.nl

Telephone: +31 24 3521348

Abstract

Numerous studies show that younger adults engage in lexically-guided perceptual learning in speech perception. Here, we investigated whether older listeners are also able to retune their phonetic category boundaries. More specifically, this research tried to answer two questions. First, do older adults show similar-sized perceptual learning effects to younger adults? Second, do differences in lexical behavior predict the strength of the perceptual learning effect? An age group comparison revealed that older listeners engage in lexically-guided perceptual learning but there were two age-related differences. Younger listeners had a stronger learning effect right after exposure than older listeners, however the effect was more stable for older than for younger listeners. Moreover, a clear link was shown to exist between individuals' lexical decision performance during exposure and the magnitude of their perceptual learning effect. A subsequent analysis on the results of the older participants revealed that, also among the older participants, with increasing age, the perceptual retuning effect becomes smaller but also more stable, mirroring the age group comparison results. These results could not be explained by differences in hearing loss.

The age effect may be accounted for by a decreased flexibility in the adjustment of phoneme categories or by age-related changes in the dynamics of spoken-word recognition, with older adults being more affected by competition from similar-sounding lexical competitors, resulting in less lexical guidance for perceptual retuning. Concluding, our results clearly show that flexibility of the speech perception system is present over the life-span.

Keywords: perceptual learning, speech perception, aging, individual differences

Introduction

Numerous studies have shown that 'ideal' listeners, i.e., young, normal-hearing, highly-educated listeners, can adapt to idiosyncratic pronunciations through lexically-guided perceptual learning in speech perception (McQueen, Cutler, & Norris, 2006; Norris, McQueen, & Cutler, 2003; see for an overview: Samuel & Kraljic, 2009), and are thus able to tune-in to a speaker to understand him/her better. The lexically-guided perceptual learning effect has been shown using a variety of exposure and test paradigms, e.g., lexical decision and phonetic categorization (e.g., Norris et al., 2003), short story presentation and phonetic categorization (e.g., Eisner & McQueen, 2006), and a picture verification procedure (e.g., McQueen, Tyler, & Cutler, in press). In the exposure phase, listeners are exposed to the idiosyncratic sound, e.g., an ambiguous sound between [s] and [f] ($/^{f}/_{s}/$), which will be learned as /s/ if heard in words such as *platypus* (because *platypus* is an existing word in English whereas *platypuf* is not), but as /f/ in words such as *giraffe* (*giraffe* is an existing word in English whereas *giras* is not). This perceptual learning effect is caused by a temporary change in phonetic category representations, rather than to changes in decision bias (Clarke-Davidson, Luce, & Sawusch, 2008). Perceptual learning has been found for tones (Mitterer, Chen, & Zhou, 2011) and different types of sounds, e.g., for stops, which differ in voice onset times (/t/ vs. /d/; Kraljic & Samuel, 2007), fricatives, which differ in noise spectra (/s/ vs. /ʃ/; Kraljic & Samuel, 2005, 2007; /s/ vs. /f/; Eisner & McQueen, 2006; Norris et al., 2003; McQueen et al., 2006; Sjerps & McQueen, 2010), liquids, which differ in liquid spectra (/l/ vs. /r/; Scharenborg, Mitterer, & McQueen, 2011), and vowels (McQueen & Mitterer, 2005).

But what about less 'ideal' listeners? Are other listener groups besides university students, e.g., young children and older listeners, also capable of perceptual learning? Or, in other words, which listener characteristics actually relate to lexically-guided perceptual learning? McQueen et al. (in press) recently showed that 6- and 12-year olds are also capable of perceptual learning, so even before children are able to read (6-year olds) they are able to use lexical knowledge to adjust phoneme categories. In the present research, we focus on flexibility of the speech perception system in an older population. More specifically, this research investigates lexically-guided perceptual learning by older listeners by comparing it to perceptual learning by younger listeners. Moreover, we investigate whether there is a link between lexical behavior during exposure and the *strength* of the perceptual learning effect. The *strength* of the perceptual learning effect can manifest itself in two ways. The *magnitude* or *size* of the perceptual learning effect is indicated by the difference between the response curves of the two exposure groups in the perceptual learning experiment. The *duration*, or *stability*, of the perceptual learning effect manifests itself as the presence (or absence, i.e., *unlearning*) of the perceptual learning effect over time. These two aspects of perceptual learning strength can be assumed to covary. If phoneme categories are less flexible, it may take more time and/or more exposure for a category change to come about. Once changed, however, it would then also take longer to undo this change, resulting in relative stability of the change. So, listeners whose phoneme categories are difficult to change are expected to show a smaller and more stable perceptual learning effect compared to listeners whose phoneme categories are more flexible.

We analyze the time-course of accepting the odd-sounding items as real words during the lexical decision exposure task for younger and older listeners. On the one hand, if the rate of acceptance is (relatively) high but does not change over exposure, this could simply be a result of a generally greater tolerance of acoustic ambiguity, while a low acceptance could be due to the ambiguous stimuli not being perceived as ambiguous. On the other hand, if there is greater acceptance of the odd-sounding items over the time-course of the lexical decision task, this suggests that items start to sound less odd. This greater acceptance over exposure trials then indicates the importance of accepting the odd-sounding items as real words for the emergence of the perceptual learning effect. Second, we investigate whether differences in frequency of acceptance of the odd-sounding items as words results in differences in the amount people shift their phoneme categories. It might be that listeners who accept more of the odd-sounding items as words during exposure will show larger category boundary shifts during testing. However, at the same time, participants could also be tolerant of ‘odd pronunciations’ during lexical decision, from the start of the experiment, and leave their categories relatively unaltered. To our knowledge, no lexically-guided perceptual learning study has ever directly investigated the link between performance during exposure and perceptual learning.

Aging may affect sensitivity to, particularly, the higher frequencies in the speech signal, which results in the loss of sensitivity to phonetic detail. This loss of sensitivity to speech detail may affect the ability to learn non-standard pronunciations, as evidence in favor of a certain pronunciation variant is weaker. Nevertheless, short-term adaptation to accents and to time-compressed speech seems to be preserved with aging and with hearing loss (Adank & Janse, 2010; Golomb, Peelle, & Wingfield, 2007; Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010; Peelle & Wingfield, 2005). Kennedy and colleagues (2009) found that aging per se did not affect the magnitude of the learning gains in perceptual skill learning: age only indirectly affected learning, via age-related declines in cognitive performance. The ability to adapt to various aspects of speech thus may remain relatively unaffected throughout one’s life. However, older adults have more language experience than younger adults, which may make their phoneme categories more resistant to change than those of younger adults. Perceptual adjustment of phoneme categories and the conditions under which these adaptations occur in an older population have not yet been investigated.

This research investigates the following questions: 1) Do older adults show similar-sized perceptual learning effects as younger adults? 2) Does lexical behavior during exposure predict the strength of the perceptual learning effect? Since we are interested in the flexibility of the speech perception system of older listeners, we aim to minimize the effect of hearing loss. We therefore use the /l/-/r/ contrast where the distinguishing information between the two consonants is mostly in those frequency regions that are supposedly to a lesser extent affected by age-related hearing loss. To ensure that the results are indeed not caused by hearing sensitivity differences, we investigate the effect of hearing loss as a control variable. The main experiment consisted of two parts (following, e.g., Norris et al., 2003 and Scharenborg et al., 2011). First, the listeners were exposed to an ambiguous [$^{l}/_{\text{ɹ}}$] in Dutch words ending on either /r/ or /l/ during a (self-paced) lexical decision task (the exposure phase). Listeners were divided into two groups: one group was exposed to the ambiguous sound only in /l/-final words and the other group was exposed to the ambiguous sound only in /r/-final words. In a subsequent (self-paced) phonetic categorization task (the test phase), listeners were confronted with a range of ambiguous sounds from the [l]-[ɹ]-continuum, appearing as the final phoneme of a non-word, and were asked to decide whether the sound was /l/ or /r/.

Method

Participants

All participants were native Dutch speakers drawn from the MPI for Psycholinguistics subject pool and they were paid for their participation. Sixteen young normal-hearing university students participated in the pretest. The ‘younger’ group consisted of 36 normal-hearing university students (8 M; mean age: 21.2), while the ‘older’ group consisted of 60 listeners aged 60+ from the Nijmegen area (24 M; mean age: 71.5; age range: 60 – 88; there was no age cut-off). Age groups differed in group size because the older adults were compared to a group of younger adults tested in a different project (Scharenborg et al., 2011). Further, investigating the effects of age and hearing loss, which were only analyzed within the older adult group, requires a larger sample size. Hearing sensitivity of the older listeners was assessed with a portable Maico ST 25 screening audiometer (air conduction thresholds only, for octave frequencies from 250 Hz – 8 kHz) in a sound-attenuated booth. A pure-tone average threshold was computed as the average over participants' thresholds at 1, 2, and 4 kHz. Six participants wore hearing aids in their daily life, but they did not wear them during testing. Mean pure-tone average (PTA, in their better ear) was 25.8 dB HL (SD=13.3; range: 5.0 – 63.3): the higher the participants' PTA, the poorer their hearing sensitivity.

Materials

For the exposure phase, 200 Dutch words were selected from CELEX (Baayen, Piepenbrock, & Gulikers, 1995). Forty words ended in /l/ and 40 ended in /r/; there were no /l/'s or /r/'s elsewhere in these 80 words. Since the sounds [l] and [ɹ] color the pronunciation of the preceding vowel, the vowel preceding [l] and [ɹ] was kept constant, such that all words ended in /əl/ or /ər/. The number of syllables was matched between the two sets of critical items (i.e., /l/-final and /r/-final words). There were 25 words with two syllables (e.g., *ezel*, donkey), 10 with three syllables (e.g., *postzegel*, stamp), and five with four syllables (e.g., *sinaasappel*, orange). Appendix A provides a full listing of all 80 critical items. Word frequency and stress patterns were matched between the two sets as far as possible.

One hundred and twenty words were selected as filler words, and 200 filler non-words were constructed. Both sets of fillers followed the same syllable-length distribution as the critical items. /l/ and /r/ did not occur in any of these items. The non-words followed Dutch phonotactic rules and tended to become non-words before their final phonemes.

All words were produced in isolation by a female native speaker of Dutch (from the Western part of the Netherlands) and digitally recorded in a sound-attenuated booth at 44 kHz. She also recorded the non-words *kwiptel* and *kwipter* for use in the test phase.

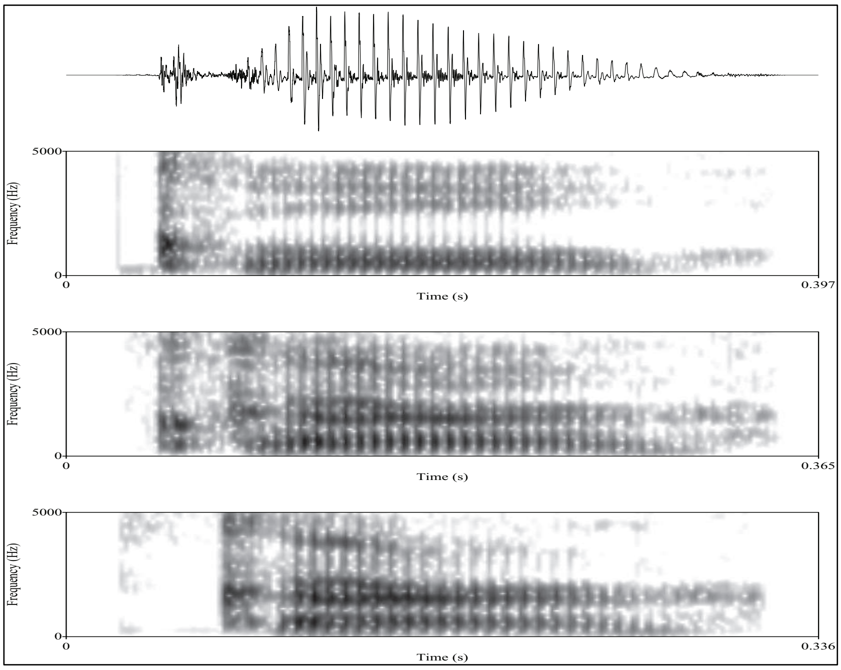

Figure 1. The top panel shows the acoustic signal for the zero-padded ambiguous syllable [kəˡ/ɹ]; the bottom 3 panels show the spectrograms of the natural versions of [kəl], the ambiguous [kəˡ/ɹ], and the natural version of [kəɹ], respectively.

Creating the Ambiguous Stimuli. From the natural recordings, versions of the 80 critical words ending in /l/ and /r/ with ambiguous final sounds were created. These ambiguous sounds [l/ɹ] and the test continuum for the phonetic categorization task were selected using a phonetic categorization pretest. The selection of the ambiguous sounds was done separately for each final syllable type present in the full set of 80 critical items for the lexical decision task. There were a total of 11 different final /Cəl/ɹ/ sequences in the set of 80 words. Note that due to devoicing of fricatives in Dutch, syllables beginning with /s/ and /z/ were treated as the same sequence, likewise for /f/ and /v/ and for /x/ and /ɣ/. The subset of words used for the pretest, i.e., one pair of words for each of the 11 sequences, is listed in Appendix B (with their English translations and non-word counterparts).

For each pair of words (e.g., *winkel* and *wekker*), the final syllable was excised using Praat (Boersma & Weenink, 2005). All excised [l]- and [ɹ]-final syllables were zero-padded at onset and offset with 25 ms of silence to allow valid pitch estimation at the start and the end of the syllable. Subsequently, each syllable received the same stylized pitch contour (based on the naturally occurring pitch contour of the final syllables in the critical items) using Praat. The resulting pairs of syllables were then each morphed to create equally-spaced 11-step continua using STRAIGHT (Kawahara, Masuda-Katsuse, & Cheveigne, 1999) in Matlab. Figure 1 shows the ambiguous syllable [kə$^{l}/_{ɹ}$] (top and third panel). This syllable was step 5 on the morphed continuum between the natural versions of [kəl] (second panel) and [kəɹ] (bottom panel). The ambiguous syllables were then concatenated, using Praat, as final syllables onto the first syllables of the matching /l/-final and /r/-final words, e.g., the morphs for /kə$^{l}/_{ɹ}$/ were concatenated with both /wɪŋ/ (yielding *winkel*) and /wɛ/ (yielding *wekker*).

The pretest stimuli were presented in three blocks, each consisting of 132 items, in a newly randomized order in each block. In each block, the sixteen pretest participants heard six [l]-[ɹ]-continuum steps (steps 1, 3, 4, 6, 7, 9), for each of the 11 syllables. These steps were chosen to sample perception of the entire continuum (excluding the endpoints). Since each morph was concatenated with both an /l/- and an /r/-final word, each morph was heard twice per block.

The task for the participants was to indicate by button press as quickly and as accurately as possible whether they heard [l] or [ɹ]. To aid the participants, the /l/-interpretation of the stimulus was shown on the bottom left of the computer screen, and the /r/-interpretation of the stimulus on the bottom right. If the /l/-interpretation was a word, the right option was a non-word, and vice versa. This procedure ensured that participants did not receive consistent lexical guidance on how to interpret the stimuli (the same ambiguous sound would receive lexical support for an /l/-interpretation on one trial, and for an /r/-interpretation on a different trial), so that phoneme category retuning would not take place. Appendix B shows the word and non-word pairs that were used. Each stimulus was presented over headphones 500 ms after trial onset. Due to an error in the testing software, the pretest for the [fəl]-[fəɹ] morphs had to be done separately. Six subjects each heard 10 repetitions of each [fəl]-[fəɹ] morph. The rest of the experimental set-up was identical to the main pretest.

The total proportions of /r/-responses to each of the tested morphs were calculated, and the most ambiguous morph was determined for each of the 11 syllables. The most ambiguous morph for syllables starting with /k, x, b/ was step 5 (step 0 is a natural [l] and step 10 a natural [ɹ]); for /m, d, f/ it was step 3; for /t, ŋ, n/ it was step 6; for /p/ it was step 2; and for /z/ it was step 7. However, after testing another separate group of six younger participants on the lexical decision task, the results showed that most of the /l/-words ending in ambiguous [$^{l}/_{ɹ}$] were not recognized as words. The ambiguous morphs were therefore changed to the next more [l]-like step for the actual experiment: /k, x, b/ = step 4; /m, d, f/ = step 2; /t, ŋ, n/ = step 5; /p/ = step 1; and /z/ = step 6. The selected morphs were then concatenated as final syllables onto the non-final syllables of the matching /l/-final and /r/-final words, as was done to create the stimuli for the pretest. This resulted in 80 stimulus pairs consisting of the same word ending in either a natural [l] or [ɹ] or the selected ambiguous [$^{l}/_{ɹ}$]. These stimuli were then used in the lexical decision task.

The stimuli used in the test phase consisted of five versions of *kwipte*$^{l}/_{ɹ}$. These were created by concatenating five different versions of the ambiguous [$^{l}/_{ɹ}$] sound as final syllables onto the first syllable *kwip* (excised from a recording of the non-word *kwipter*). The steps (i.e., steps 2, 4, 5, 6, and 8; where step 5 was judged to be the most ambiguous sound in the pretest) were taken from the [təl]-[təɹ]-continuum created for the pretest. Both the /l/-final and /r/-final reading of the resulting string is a non-word in Dutch.

Procedure

Two experimental-word lists were created in which the test items appeared in a pseudo-randomized running order. The restrictions were that no critical item (i.e., no word ending in [$^{l}/_{ɹ}$]) was allowed to appear in the first six words, and no two critical items could appear within a range of four words. Each list consisted of 400 words, i.e., the 200 non-words, 120 filler words, 40 words ending in a natural [l] or [ɹ], and 40 critical items, i.e., the /r/-final or /l/-final words ending in [$^{l}/_{ɹ}$]. The difference between the two word lists was that one list contained only natural /r/-final words and /l/-final words ending in [$^{l}/_{ɹ}$], the other list contained the natural /l/-final words and the /r/-words ending in [$^{l}/_{ɹ}$]. The younger and older listener groups were split into two groups and assigned one of the two experimental-word lists. Eighteen younger and 30 older listeners took part in the group who heard ambiguous /l/-final words during exposure and 18 younger and 30 older listeners took part in the group who heard ambiguous /r/-final words during exposure.

Participants were tested individually in a sound-treated booth. The stimuli were presented binaurally over closed headphones. Participants were asked to press a button as fast and accurately as possible when they heard a word (left button) or a non-word (right button). They were not informed about the presence of ambiguous sounds. The lexical decision task lasted approximately 25 minutes. Subsequently, participants were tested using a phoneme categorization test. They were asked to decide as fast and accurately as possible, by button press, whether the stimulus ended in /l/ or in /r/. The five ambiguous *kwipte*$^{l}/_{ɹ}$ stimuli were each presented six times per block, and were newly randomized for each of a total of three blocks (90 items in total). The /l/-interpretation of the stimulus (*kwiptel*) was shown on the bottom left of the computer screen and the /r/-interpretation of the stimulus on the bottom right (*kwipter*). The phonetic categorization task lasted approximately 8 minutes. All stimuli for all participants were presented at an average intensity level of 75 dB SPL.

Results

We investigate whether older listeners show a similar-sized perceptual effect as younger listeners through an age-group comparison. To that end, we investigate performance and the time-course of accepting ambiguous items as words in the lexical decision task and the perceptual learning effect as exhibited in the phonetic categorization task. The final analyses investigate whether age (and hearing sensitivity) predicts perceptual learning.

All analyses were carried out using generalized linear mixed-effect models (e.g., Baayen, Davidson, & Bates, 2008), containing both fixed and random effects, using the logit link function. The fixed and random factors differed per analysis and are therefore listed for each analysis separately. The parameters of the generalized linear models are set using maximum likelihood estimation. We used dummy coding. A generalized model has the form

$$\text{logit}\ p = c + \beta_1 Factor_1 + \beta_2 Factor_2 + \beta_3 Factor_3 + \ldots,$$

where logit *p* represents log [*p*(1-*p*)]; logit *p* is the ‘dependent variable’. The constant *c* is the intercept. The different *β*s (Chatterjee, Hadi, & Price, 2000) represent the relevance (effect size) of the different predictors for the estimation of the logit *p*. In each analysis, a best-fitting model is built using the fixed and random variables. We started by building the most complex model, i.e., the model with all possible interactions between the predictors. Subsequently, interactions and predictors that proved not significant were step-by-step removed from the model. The best-fitting model only contains predictor variables and interactions that are significant. We only report statistically significant effects and the absolute estimated values of the different *β*s, with an explanation of the found effect.

Lexical Decision

In accordance with Norris et al. (2003), participants who judged fewer than 20/40 of the [$^{l}/_{ɹ}$]-items as words were excluded from further analyses due to ‘poor’ performance. This resulted in the exclusion of one participant of the younger group, who heard natural /l/-final words and ambiguous /r/-final words, and nine participants of the older group (none of whom wore hearing aids in daily life): four participants who heard natural /r/-final and ambiguous /l/-final words and five participants who heard natural /l/-final and ambiguous /r/-final words, leaving 51 older listeners.

Percentage of ‘yes’ responses for the non-word filler items was 1.8% for the younger listeners and 2.2% for the older listeners. Table 1 shows the mean percentages of ‘yes’ responses for the natural and the ambiguous versions of the /l/-final and /r/-final words for the listeners who were exposed to the ambiguous sound in /r/-final or /l/-final words. Listeners from both age groups accepted most stimuli ending in [$^{l}/_{ɹ}$] as words. Note that in many cases, listeners did not need the final phoneme to recognize the word: many of the multi-syllabic words used in the present study become unique before the final phoneme.

Table 1

Performance on the Lexical Decision Task, Mean Percentage of ‘Yes’ Responses for the Natural and the Ambiguous Versions of the /l/- and /r/-final Words

	Natural liquids		Ambiguous liquids	
	/l/-final	/r/-final	/l/-final	/r/-final
Younger: Mean % ‘yes’	98.2	99.0	92.6	85.9
Older: Mean % ‘yes’	97.0	97.4	91.6	87.9

The responses to the natural, filler words, and ambiguous items were subsequently analyzed statistically. We investigated the question whether listeners immediately accept ambiguous items as words or whether they get more tolerant over trials. The critical issue here is the time course of accepting an ambiguous target stimulus as a word during the lexical decision task. We therefore only focus on the items that could potentially elicit a 'yes it is a word response', thus ignoring the nonword items. The responses of the two 'exposure' groups were taken together and two new categories were created: 'natural', containing the responses to the natural items in the two exposure groups (i.e., the natural /l/-final words in the group of listeners who learned to map the ambiguous sound onto /r/ and the natural /r/-final words in the group of listeners who learned to map the ambiguous sound onto /l/), and 'ambiguous', containing the responses to the ambiguous items in the two exposure groups. The third type of words were the word fillers (on the intercept). The dependent variable was whether the response to the word type was 'yes' (coded as 1) or 'no' (coded as 0). Fixed predictors were trial, word type (natural, ambiguous, or word filler), and age group (younger vs. older, the former group is on the intercept). Item and subject were the random predictors.

Fewer 'yes it is a word' responses were given to ambiguous items (89.6%; $\beta = -2.9283$, $SE = .4325$, $p < .001$) than to the filler words (97.1%) by both age groups. However, over trials, listeners from both age groups started to give more 'yes' responses to the ambiguous items ($\beta = .0064$, $SE = .0017$, $p < .001$), so both the younger and the older listeners seemed to 'learn' to accept the ambiguous items over the course of the experiment. Moreover, older listeners gave fewer 'yes' responses to the natural items (97.2%; $\beta = -.7978$, $SE = .3183$, $p < .05$) than the younger listeners (98.6%). Fewer 'yes' responses were given to the word fillers by the younger listeners over trials ($\beta = -.0023$, $SE = .0009$, $p < .01$). Since there is no significant difference between the number of 'yes' responses to word fillers and natural words, we can conclude that younger listeners seem to get less sure over trials about whether the natural items were words or not. An interaction between trial and age group showed that this growing uncertainty over trials was less for the older listeners (β = .0020, *SE* = .0008, $p < .01$). Perhaps the task itself made the younger listeners more cautious over trials, making them more uncertain on the natural items as well.

Summarizing, the results showed that both the younger and older listeners who were exposed to the ambiguous [$^{l}/_{\text{ɹ}}$] in the normally /l/-final words tended to interpret the ambiguous sound as /l/, whereas listeners who were exposed to [$^{l}/_{\text{ɹ}}$] in the context of normally /r/-final words interpreted [$^{l}/_{\text{ɹ}}$] as /r/. Moreover, younger and older listeners showed a similar time-course effect on the lexical decision task. Listeners get more tolerant over trials in that they accept more ambiguous items as words over the course of the lexical decision task.

Phonetic Categorization

Figure 2 shows the proportion of /l/- and /r/-responses for the five ambiguous stimuli in the phonetic categorization task for the three blocks, separately. The responses for the listeners who were exposed to [$^{l}/_{\text{ɹ}}$] in the normally /r/-final words are indicated with ‘r’s for the younger listeners and with ‘R’s for the older listeners. The responses for the listeners who were exposed to [$^{l}/_{\text{ɹ}}$] only in the normally /l/-final words are indicated with ‘l’s for the younger listeners and with ‘L’s for the older listeners. The responses to the five ambiguous stimuli were subsequently analyzed (the dependent variable is whether the response is /l/, coded as 0, or /r/, coded as 1). The fixed predictors were exposure group (exposed to the ambiguous sound only in /l/-final words (on the intercept) or only in /r/-final words during the lexical decision task), age group (older listener group is on the intercept), and test block; stimulus step (stimulus step is a continuous variable with step 3 on the intercept, the steps are not linearly spaced) was used as a control variable. Subject was the only random factor as all ambiguous sounds were embedded in the same *kwipte*^{l}/ɹ nonword context. In our report of the analysis, we focus on those results that are relevant to the research question.

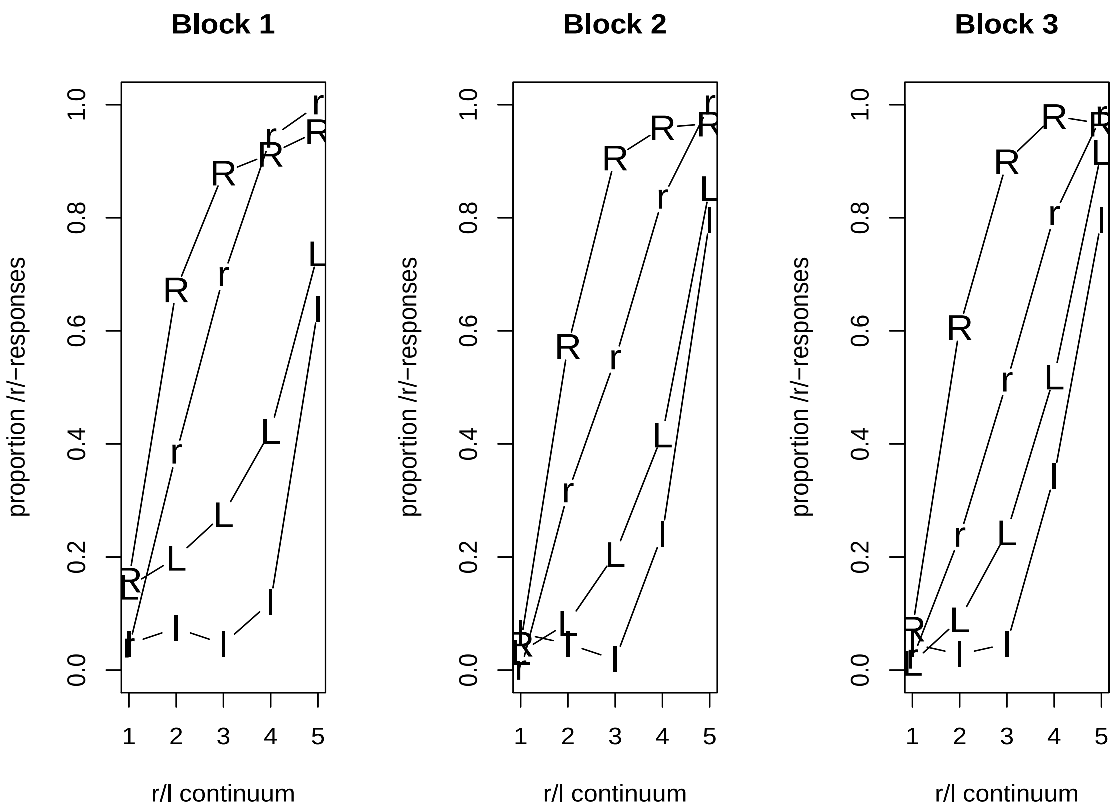

Figure 2. The total proportion of /r/ responses for the two exposure groups per test block: r and R indicate the groups of younger and older listeners, respectively, who learned to map [^{l}/ɹ] onto [ɹ]. l and L indicate the groups of younger and older listeners, who learned to map [^{l}/ɹ] onto [l] for the five ambiguous test stimuli.

Table 2 displays the parameter estimates in the best-fitting model of performance. Both age groups show an effect of exposure group on phonetic categorization. In general, older listeners who were exposed to [^{l}/ɹ] in the normally /r/-final words were strongly biased to label the sounds on the continuum as /r/, while those older listeners who were exposed to [^{l}/ɹ] in the normally /l/-final words were less likely to do so. This difference between the curves of the /r/-responses for the two exposure groups (within an age group) is the perceptual learning effect. The magnitude of the difference between the mean proportion of /r/ responses made by listeners in the /r/-final exposure group and /l/-final group was stronger for younger listeners in block 1 (18.5% vs. 61.2%, for the /l/- vs. /r/-exposure group, respectively) than for older listeners (35.6% vs. 71.6%, for the /l/- vs. /r/-exposure group, respectively; shown by the exposure group and age group interaction), i.e., the younger listeners not only also showed a perceptual learning effect, the magnitude of the perceptual learning effect is significantly *larger* for the younger listeners than for the older listeners in block 1. The difference in magnitude in perceptual learning between the age groups decreases in subsequent blocks, as witnessed by a three-way interaction between block, listener group, and exposure group (note that if we would average the results over all test blocks, the size of the perceptual learning effect between the two age groups does not differ). Table 3 shows the size of the perceptual learning effect per age group. Indeed, a per block analysis showed that the interaction between age group and exposure group is no longer significant in the later test blocks (Block 2: β = -.5216, SE = .9369, $p > 0.5$; Block 3: β = -1.4653, SE = .9882, $p > .1$). The younger listeners thus showed ‘unlearning’, while the learning effect for the older listeners remained stable over blocks[1].

[1] A separate experiment in which two new groups of listeners (18 younger listeners and 40 older listeners, aged 60+) were tested on the phonetic categorization task only (thus without the exposure phase) showed that the initially larger learning effect and unlearning over blocks by the younger adults could not be explained by a potential /r/-bias for the older listeners.

Table 2

Lexically-guided Perceptual Learning: Fixed Effect Estimates for the Best-fitting Model of Performance in the Phonetic Categorization Task (n = 7740)

Fixed effect	*β*	Standard Error	*p*<
Intercept	-1.1256	.3483	.01
Exposure group	2.8760	.5032	.001
Block	-.3162	.0862	.001
Age group	-1.8580	.5687	.01
Stimulus step	.7370	.0556	.001
Block × Exposure group	.4835	.1359	.001
Block × Age group	.4997	.1589	.01
Exposure group × Age group	1.7117	.8130	.05
Stimulus step × Age group	.3697	.1057	.001
Stimulus step × Exposure group	.4192	.0913	.001
Stimulus step × Block	.4984	.0550	.001
Block × Age group × Exposure group	-1.4666	.2375	.001
Stimulus step × Age group × Exposure group	.9548	.1781	.001
Stimulus step × Block × Exposure group	-.2606	.0766	.001
Stimulus step × Block × Age group	-.3185	.0799	.001

Summarizing, both the younger and older listener groups show perceptual learning of ambiguous sounds on the /l/-/r/ continuum. However, returning to our first research question:

the perceptual learning effect is larger right after exposure for younger listeners, while the effect is more stable for older listeners[2].

Table 3

Difference in Mean Proportion of /r/-responses by the Listeners in the /r/-final Exposure Group versus the /l/-final Exposure Group for the Younger and Older Listeners

	Test Block 1	Test Block 2	Test Block 3
Younger listeners	42.7	30.6	25.8
Older listeners	36.0	37.3	34.9

Predicting Phonetic Categorization Performance from Lexical Decision Performance

The following analyses investigated whether lexical behavior during exposure and age differences (among the older adults) predict the strength of the perceptual learning effect. The first analysis investigates whether differences in frequency of acceptance of the odd-sounding items as words during the lexical decision task results in differences in the amount people shift their phoneme categories. To that end, we investigated whether listeners who more often judged ambiguous items to be words during the lexical decision task gave more learning-consistent responses (i.e., more /r/-responses when exposed to the ambiguous sound in /r/-final words and more /l/-responses when exposed to the ambiguous sound in /l/-final words). We focus on the *ambiguous* stimuli, because these are the crucial items in the lexical decision task that are supposed to induce phonetic category adjustment. Per age group, the two

[2] We also ran the analyses without removing any subjects (thus including the 10 listeners who had been excluded on the basis of the too few ‘yes’ responses to the ambiguous items during exposure). The most important age-related conclusions, i.e., an initially larger learning effect for younger listeners and a more stable learning effect for older listeners, were still found.

exposure groups were taken together and a new category ‘learning-consistent’ was created in which the /r/-responses during the phonetic categorization task of the group of listeners exposed to the ambiguous sound in /r/-final words and the /l/-responses of the group of listeners exposed to the ambiguous sound in the /l/-final words were combined. Moreover, we only analyzed data from the stimulus steps of interest (i.e., the most ambiguous steps, steps 2, 3, and 4). Percentages of ambiguous items accepted as words during the lexical decision task were calculated for each participant, and used as a fixed predictor of whether the category response was learning consistent (the dependent variable, coded as 0 and 1, for not learning consistent and learning consistent, respectively). Age group (younger listener group on the intercept), stimulus step, and test block were used as control variables; subject was the random factor.

Table 4

Fixed Effect Estimates for the Best-fitting Model of 'Learning-Consistent' Performance in the Phonetic Categorization Task (n = 4644)

Fixed effect	β	Standard Error	$p<$
Intercept	2.8420	.3267	.001
Acceptance percentage of ambiguous words	7.0385	1.6190	.001
Age group	-1.3091	.4114	.005
Block	-.5290	.0880	.001
Block × Age group	.5705	.1086	.001
Stimulus step	1.6113	.2466	.001
Stimulus step × Age group	-1.4805	.2982	.001
Stimulus step × Block	-.4074	.1068	.001
Stimulus step × Block × Age group	.3333	.1323	.05

Table 4 displays the parameter estimates in the final model. Listeners who more often judged an ambiguous item as a word in the lexical decision task gave more learning-consistent responses during the phonetic categorization task, i.e., they showed stronger perceptual learning than listeners who less often judged an ambiguous item as a word. These listeners thus seem to have retuned their phoneme categories more.

Subsequently, the effect of age on the perceptual learning effect was investigated among the *older* participants. Since hearing loss is a common phenomenon among older listeners, hearing loss was used as a control variable (centralized to the mean). The dependent variable was again the category ‘learning-consistent’, taking into account only the three most ambiguous stimulus steps. Age and hearing sensitivity were shown to be correlated (r = .40, p < .005). To reduce collinearity in the model, a residual was created for age (with hearing loss partialed out), which was used as a fixed predictor. Other control variables were stimulus step and test block. Subject and item were the random factors.

Consistent with the age-group comparison where the younger listeners initially had a larger perceptual learning effect than the older listeners, the perceptual learning effect in the first block was smaller with increasing age (β = -.1142, SE = .0385, p < .005). Moreover, the decrease of the perceptual learning effect over blocks was smaller with increasing age (β = .04473, SE = .0133, p < .001), which is also consistent with the results of the group comparison. Hearing sensitivity does not modify the size or stability of the perceptual learning effect, showing that our stimuli are indeed audible for listeners with hearing loss.

General discussion

Our research was inspired by numerous findings that young university students show lexically-guided perceptual learning (for an overview: Samuel & Kraljic, 2009). We focus on flexibility of the speech perception system in an *older* population. More specifically, this research tries to answer the following two questions. Do older adults show similar-sized perceptual learning effects as younger adults? This was investigated by comparing lexically-guided perceptual learning of older listeners (aged 60+) to that of younger listeners. The second question asked whether lexical behavior during exposure predicts the strength of the perceptual learning effect. To ensure that the results are not caused by age differences in hearing sensitivity, the effect of hearing loss is investigated as a control variable.

The perceptual learning experiment consisted of two parts: an exposure phase consisting of a lexical decision task and a phonetic categorization test phase. The lexical decision results showed that both the younger and older listeners interpreted the majority of the stimuli with ambiguous final [l/ɹ] as words. Moreover, the time course of accepting the ambiguous items as words during exposure is similar for both age groups: listeners in both age groups show increased acceptance of ambiguous items as words over the course of the lexical decision task. This finding shows how increased acceptance of the odd-sounding items as real words may reflect the perceptual learning effect.

Despite the similarity in time course of acceptance of ambiguous items over exposure, there were age-related differences in the strength of the perceptual learning effect: the learning effect was stronger right after exposure for the younger listeners, but the effect was more stable for older listeners, i.e., younger listeners showed 'unlearning' whereas older listeners did not. This age effect was confirmed in a subsequent analysis of the effect of age among the *older* adults; also within the group of older listeners, the perceptual learning effect became smaller but also more stable with increasing age. Importantly, this different pattern of perceptual learning could not be explained by age-related differences in hearing loss, since no effect of hearing loss on the strength of the perceptual learning effect was found.

These findings raise the question what 'age' is or what it represents. As in Kennedy and colleagues (2009), age per se might not have affected perceptual learning directly, but mainly indirectly via age-related changes in cognitive or linguistic abilities that we did not explicitly measure. Older persons, for instance, have (much) more linguistic experience than younger persons. Perhaps language experience makes phonetic categories more robust and resistant to larger or faster changes for older adults, while younger adults may have sparser, more malleable categories. As argued in the Introduction, if category changes take more time and/or need more compelling evidence, undoing these changes would also take longer, resulting in relatively stable learning. Even though we did not find an age group difference of acceptance of the ambiguous words as words over all exposure trials, a subset analysis on the first two thirds of the 200 (ambiguous, natural, and filler) word exposure trials showed indeed a stronger initial reluctance in accepting the ambiguous stimuli as words by the older listeners compared to the younger listeners (β = -.0180, *SE* = .0081, $p < .05$). This age difference in initial acceptance mirrors the age difference in phoneme category flexibility observed in the phonetic categorization task. Note, however, that the smaller but more stable perceptual retuning effect with increasing age was also found *among* the older adults. As one might expect linguistic experience to plateau in older age, this linguistic experience account may not provide a full explanation.

Another explanation for the age difference in category adjustment could be an age-related decline in efficiency of inhibitory processes (Hasher & Zacks, 1988; Zacks & Hasher, 1994; Mattys & Scharenborg, under review). Such a reduction of efficient inhibitory processes might affect the dynamics of spoken-word recognition by resulting in less de-activation of similar-sounding lexical candidates in older adults. Older listeners have indeed been shown to be more affected by competition from similar-sounding words than younger listeners (cf. Ben-David, Chambers, Daneman, Pichora-Fuller, Reingold, & Schneider, 2011; Sommers, 1996; Sommers & Danielson, 1998). We conjecture that by keeping more word candidates activated during the word recognition process, lexical guidance from the critical words may be less compelling, resulting in decreased lexically-induced perceptual learning.

In addition to age, lexical behavior during the lexical decision task also predicts the strength of the perceptual learning effect. Listeners who more often gave 'yes' responses to ambiguous items during the lexical decision task showed stronger perceptual learning. In other words, people vary in the amount they shift their phoneme categories on the basis of lexical guidance. This provides evidence that it is generally *not* the case that participants are tolerant during exposure (accepting odd-sounding items as words), and yet leave their category boundaries unaltered. As far as we are aware, this is the first time that a link between the frequency of 'yes' responses to ambiguous items during the lexical decision task and the strength of the perceptual learning effect is shown.

The unlearning effect for younger listeners was also reported by Mitterer et al. (2011) for the lexically-guided perceptual learning of tones. Our results on perceptual learning of a consonant contrast with distributed (i.e., non-local) acoustic cues (as the consonant also affects the quality of the preceding vowel) and Mitterer et al.'s results on perceptual learning of tones, whose cues are also non-local, seems to suggest that perceptual learning of a contrast with non-local cues differs from that of plosive or fricative contrasts as used in other studies which differ primarily in local acoustic cues. An explanation for why such differences between local and non-local cues would impact on the stability of the learning effect is however lacking.

To conclude, older listeners, like younger listeners, show perceptual learning of a liquid contrast. Together with the results found by McQueen et al. (in press) on 6-year and 12-year olds, these data clearly show that the ability of lexically-driven perceptual learning is present over the life-span. Nevertheless, an age-related decline in the size of the perceptual learning effect and an increase in its stability were observed, which may be accounted for by a decreased flexibility in the adjustment of phoneme categories or by age-related changes in the dynamics of spoken-word recognition.

References

Adank, P. & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and Aging, 25(3)*, 736-740.

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.

Baayen, R., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (release 2).* PA: Linguistic Data Consortium, University of Pennsylvania.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., & Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements. *Journal of Speech, Language, and Hearing Research*, *54*, 243-262.

Boersma, P. & Weenink, D. (2005). *Praat. Doing phonetics by computer (Version 5.1).*

Chatterjee, S., Hadi, A. S., Price, B. (2000). *Regression analysis by example*. New York: John Wiley & Sons.

Clarke-Davidson, C. M., Luce, P. A., & Sawusch, J. R. (2008). Does perceptual learning in speech reflect changes in phonetic category representation or decision bias? *Perception & Psychophysics*, *70(4)*, 604-618.

Cutler, A. (under review). Native listening: Can it be diagnosed?

Eisner F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America*, *119*, 1950-1953.

Goldstone, R. L. (1998). Perceptual learning. *Annu. Rev. Psychol.*, *49*, 585-612.

Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *Journal of the Acoustical Society of America*, *121*, 1701-1708.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Schurman, J. (2010). Short-term adaptation to accented English by younger and older adults. *Journal of the Acoustical Society of America Express Letters*, *128(4)*, 200-204.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Janse, E., & Adank, P. (in press). Predicting foreign-accent adaptation in older adults. Manuscript submitted to *Quarterly Journal of Experimental Psychology*.

Kawahara, H., Masuda-Katsuse, I., & Cheveigne, A. (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: possible role of a repetitive structure in sounds. *Speech Communication*, *27*, 187-207.

Kennedy, K. M., Rodrigue, K. M., Head, D., Gunning-Dixon, F., & Raz, N. (2009). Neuroanatomical and cognitive mediators of age-related differences in perceptual priming and learning. *Neuropsychology*, *23*, 475-491.

Kraljic T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*, 141-178.

Kraljic T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*, 1-15.

Mattys, S. L., & Scharenborg, O. (under review). Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30(6)*, 1113-1126.

McQueen, J. M., Tyler, M., & Cutler, A. (in press). Lexical retuning of children's speech perception: Evidence for knowledge about words' component sounds. *Language Learning and Development.*

McQueen, J. & Mitterer, H. (2005). Lexically-driven perceptual adjustments of vowel categories. *Proceedings of ISCA Workshop on Plasticity in Speech Perception* (p. 233-236). Geneva, France: ISCA.

Mitterer, H., Chen, Y., & Zhou, X. L. (2011). Phonological Abstraction in Processing Lexical-Tone Variation: Evidence From a Learning Paradigm. *Cognitive Science, 35*, 184-197. doi:10.1111/j.1551-6709.2010.01140

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47(2)*, 204-238.

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology: Human Perception and Performance*, *31(6)*, 1315-1330.

Samuel, A. G., & Kraljic, T. (2009) Perceptual learning in speech perception. *Attention, Perception & Psychophysics*, 71, 1207-1218.

Scharenborg, O., Janse, E., & Weber, A. (2012). Perceptual learning of /f/-/s/ by older listeners. To appear in *Proceedings of Interspeech 2012*, Portland, OR.

Scharenborg, O., Mitterer, H., & McQueen, J. M. (2011). Perceptual learning of liquids. *Proceedings of Interspeech*, Florence, Italy.

Sjerps, M. J., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 195-211.

Sommers, M. S. (1996). The structural organization of the mental lexicon and its contribution to age-related changes in spoken word recognition. *Psychology and Aging, 11*, 333–341.

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging, 14*, 458–472.

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory and language* (pp. 241-264), San Diego, CA: Academic Press.

Appendix A

Overview of the 40 /l/-final and 40 /r/-final words used in the lexical decision task.

/l/-final words			/r/-final words		
2-syllables	3-syllables	4-syllables	2-syllables	3-syllables	4-syllables
appel	amandel	acceptabel	bakker	aansteker	apotheker
bijbel	beginsel	detailhandel	cijfer	bewaker	medestander
deksel	bindmiddel	geneesmiddel	danser	huishoudster	misdadiger
ezel	boekhandel	sinaasappel	emmer	kabouter	psychiater
fakkel	gemompel	voedingsmiddel	fietser	ontvanger	wetenschapper
handel	obstakel		honger	onzeker	
hemel	onnozel		kapper	schoenmaker	
heuvel	pantoffel		kijker	wethouder	
kachel	postzegel		kikker	wijsvinger	
kapsel	wastafel		masker	zeewater	
knuppel			meter		
kogel			modder		
mantel			oever		
meubel			peper		
mossel			puber		
nagel			spijker		
schotel			steiger		
snavel			suiker		
spiegel			tijger		
stempel			venster		
stengel			veter		

tegel	vijver
tunnel	wekker
winkel	winter
zadel	zender

Appendix B

Overview of the word and non-word pairs used in the pretest.

/l/-final			/r/-final		
word	English translation	non-word	word	English translation	non-word
stengel	stalk	*stenger*	*honger*	hunger	*hongel*
amandel	almond	*amander*	*zender*	channel	*zendel*
tunnel	tunnel	*tunner*	*doener*	doer	*doenel*
meubel	furniture	*meuber*	*puber*	teenager	*pubel*
tegel	tile	*teger*	*tijger*	tiger	*tijgel*
heuvel	hill	*heuver*	*oever*	shore	*oevel*
winkel	shop	*winker*	*wekker*	alarm-clock	*wekkel*
hemel	heaven	*hemer*	*emmer*	bucket	*emmel*
ezel	donkey	*ezer*	*danser*	dancer	*dansel*
schotel	dish	*schoter*	*veter*	shoelace	*vetel*
stempel	stamp/seal	*stemper*	*kapper*	hairdresser	*kappel*