

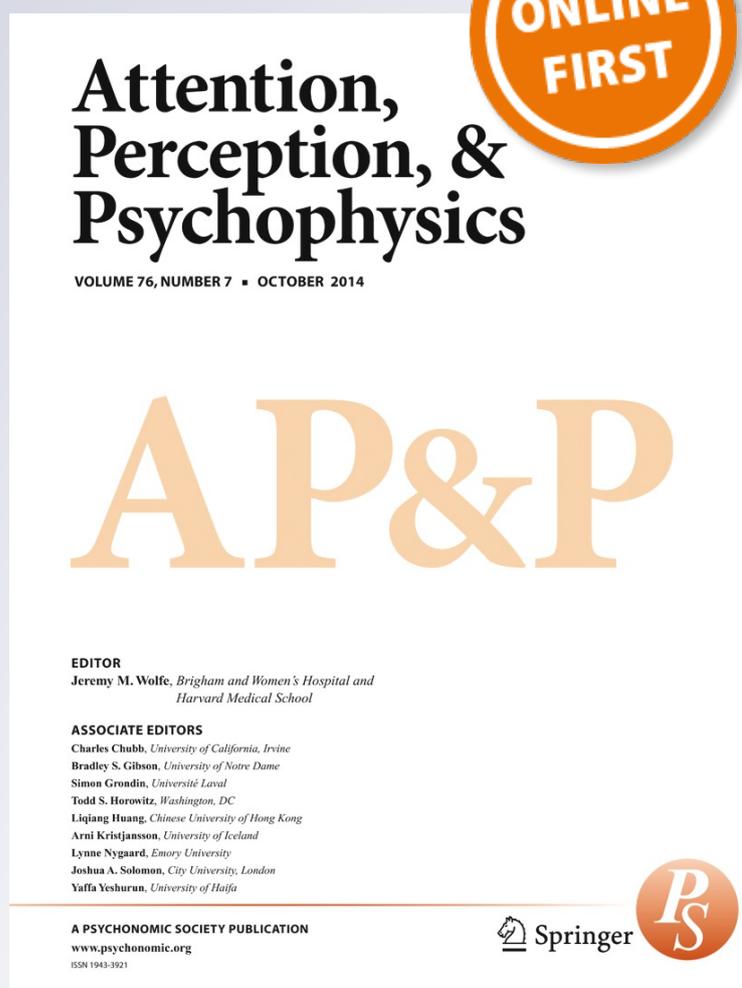
The role of attentional abilities in lexically guided perceptual learning by older listeners

Odette Scharenborg, Andrea Weber & Esther Janse

Attention, Perception, & Psychophysics

ISSN 1943-3921

Atten Percept Psychophys
DOI 10.3758/s13414-014-0792-2



Your article is protected by copyright and all rights are held exclusively by The Psychonomic Society, Inc.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

The role of attentional abilities in lexically guided perceptual learning by older listeners

Odette Scharenborg · Andrea Weber · Esther Janse

© The Psychonomic Society, Inc. 2014

Abstract This study investigates two variables that may modify lexically guided perceptual learning: individual hearing sensitivity and attentional abilities. Older Dutch listeners (aged 60+ years, varying from good hearing to mild-to-moderate high-frequency hearing loss) were tested on a lexically guided perceptual learning task using the contrast [f]-[s]. This contrast mainly differentiates between the two consonants in the higher frequencies, and thus is supposedly challenging for listeners with hearing loss. The analyses showed that older listeners generally engage in lexically guided perceptual learning. Hearing loss and selective attention did not modify perceptual learning in our participant sample, while attention-switching control did: listeners with poorer attention-switching control showed a stronger perceptual learning effect. We postulate that listeners with better attention-switching control may, in general, rely more strongly on bottom-up acoustic information compared to listeners with poorer attention-switching control, making them in turn less susceptible to lexically guided perceptual learning. Our results, moreover, clearly show that lexically guided perceptual learning is not lost when acoustic processing is less accurate.

Keywords Perceptual learning · Speech perception · Attention · Aging · Individual differences

O. Scharenborg (✉) · E. Janse
Centre for Language Studies, Radboud University Nijmegen,
Erasmusplein 1, 6525 HT Nijmegen, The Netherlands
e-mail: O.Scharenborg@let.ru.nl

O. Scharenborg · A. Weber · E. Janse
Donders Institute for Brain, Cognition, and Behaviour, Radboud
University Nijmegen, Nijmegen, The Netherlands

A. Weber · E. Janse
Max Planck Institute for Psycholinguistics, Nijmegen, The
Netherlands

Using the perceptual learning paradigm, ample evidence has been gathered by now showing that young listeners use lexical and phonotactic knowledge to quickly retune their phonemic categories in response to ambiguous pronunciations of sounds (e.g., Cutler, McQueen, Butterfield, & Norris, 2008; Norris, McQueen, & Cutler, 2003; see for an overview Samuel & Kraljic, 2009). Ultimately, listeners are listening for meaning, and they rather interpret words with ambiguous sounds as meaningful words than nonsense words. In this sense, the lexicon guides listeners' adjustment of phonemic categories when they encounter ambiguously produced sounds. The quick retuning of phonemic categories helps listeners to understand new speakers and unfamiliar accents as it allows them to easily comprehend other words produced by those speakers (Norris et al., 2003).

In a typical lexically guided perceptual learning experiment, listeners are first exposed to an ambiguous sound, e.g., an ambiguous sound between [f] and [s] ([f/s]) in a word (Norris et al., 2003). During this exposure phase, in our study a lexical decision task, the ambiguous sound [f/s] will be learned to be interpreted as /s/ if heard in words such as *platypus* (*platypuf* is a non-word in English), or as /f/ in words such as *giraffe* (*giras* is a non-word in English). This learning generalizes to words that have not been presented earlier (McQueen, Cutler, & Norris, 2006), so that English young adults interpret the previously unheard word [na f/s], as *nice* or *knife* depending on their previous exposure to *platypu[f/s]* or *gira[f/s]*, respectively (see also Davis et al., 2005 and Hervais-Adelman et al., 2008, on the role of lexicality in perceptual learning of noise-vocoded speech). Recently, it has been shown that lexically guided perceptual learning is presumably present over a person's life span. McQueen, Tyler, and Cutler (2012) showed that 6- and 12-year olds are already capable of perceptual learning. Furthermore, Scharenborg and Janse (2013) found that older Dutch listeners (60+ years old) are also capable of perceptual learning and show a lexically

guided perceptual learning effect comparable to that of younger listeners on a liquid contrast. The ability to adapt to new listening conditions thus remains present throughout a life span (see also, e.g., Adank & Janse, 2010; Golomb, Peelle, & Wingfield, 2007; Peelle & Wingfield, 2005). In the present paper, we further investigate the mechanisms that enable this category retuning process with a special focus on the hearing and attentional abilities of listeners.

Lexically guided perceptual learning has been found to be stronger the more often a listener considers words with an ambiguous sound as real words during exposure (Scharenborg & Janse, 2013); this suggests that there are differences in the amount of lexical guidance and subsequent category retuning between listeners. Crucially, the weight people assign to the lexicon versus acoustic detail in the speech signal seems to vary with the difficulty of the listening conditions. In a series of experiments in which the use of lexical information and acoustic detail were pitted against one another, Mattys and colleagues (2009) investigated this weighting of lexical information versus acoustic detail in the presence of noise and cognitive load. In their study, listeners were presented with ambiguous two-word stimuli that varied on a boundary continuum from *mild#option* (# indicates a word boundary, both parts of the stimulus are English words) to *mile#dooption* (word/nonword combination). When the speech signal contained severe background noise, listeners relied relatively more on the acoustic detail that can be glimpsed from the target speech even if this led to lexically incorrect segmentations (Mattys, Brooks, & Cooke, 2009). That is, listeners relatively more often segmented the ambiguous two-word stimuli into the word/non-word combination in severe noise compared to a no-noise condition. On the other hand, listeners were found relatively more often to segment them into two existing words in the presence of a cognitive load due to a dual task compared to a single-task condition (Mattys et al., 2009; Mattys & Wiget, 2011). Stronger reliance on lexical knowledge in the presence of increased cognitive load could either be driven by attention being drawn away from acoustic detail or by a decreased ability of listeners to inhibit lexical activation. Both will cause listeners to fall back on lexical knowledge (Mattys, Seymour, Attwood, & Munafò, 2013). If taking away attentional resources makes listeners rely more strongly on lexical information, this raises the question whether there are individual differences in the size of the lexically guided perceptual learning effect depending on one's cognitive abilities, and attentional abilities in particular.

For lexically guided perceptual learning to occur, it is critical that listeners can perceive the sound ambiguity and that they can rely on lexical knowledge for retuning a sound category. This study investigated two possible variables that could therefore influence lexically guided perceptual learning: (1) hearing sensitivity; i.e., how well can listeners hear the contrast and thus the ambiguity; (2) attentional abilities; i.e., to

what extent are listeners able to use lexical content for interpreting deviant pronunciations. We investigated these potential driving forces of lexically guided perceptual learning among older listeners, since older listeners naturally vary more in hearing and attentional abilities than younger listeners, university students in particular. This variability is needed in order to avoid ceiling performance (i.e., everyone doing very well) and to be able to observe differences between listeners' hearing and cognitive abilities and the amount of lexically guided perceptual learning.

Aging often affects sensitivity to the higher frequencies in speech. Information in these higher frequencies is thus (presumably) less audible for many older listeners. The sound contrast used in this study therefore consisted of a contrast which has its distinguishing information mainly in the higher frequencies. We chose a fricative contrast, since fricatives have been found to frequently induce recognition errors for normal-hearing listeners and particularly for listeners with (simulated) sensorineural hearing loss (e.g., Maniwa, Jongman, & Wade, 2008; Sher & Owens, 1974; Zeng & Turner, 1990). Specifically we chose the /s/-/f/ contrast, a contrast that has also been used in many other studies on lexically guided perceptual learning (e.g., Eisner & McQueen, 2005; Norris et al., 2003; McQueen et al., 2006). [s] has most of its energy concentrating around 5000 Hz, while [f] has a flatter spectrum with the energy distributed more uniformly over the spectrum (Rietveld & van Heuven, 1997).

Reduced sensitivity to the higher frequencies results in reduced sensitivity to phonetic detail in speech. This has been shown to have an impact on the recognition of accented speech (Adank & Janse, 2010; Gordon-Salant, Yeni-Komshian, & Fitzgibbons, 2010a, b; Janse & Adank, 2012). The particular effect of high-frequency hearing loss on lexically guided perceptual learning is however unclear. Considering the importance of positive perceptual evidence of the ambiguity of a sound and the impact of age-related hearing loss on the identification of speech sounds, we expected hearing loss to interfere with perceptual learning as perceptual evidence in favour of a certain pronunciation variant should be weaker. Furthermore, increased perceptual effort in decoding the auditory stimuli due to hearing loss has been shown to take away resources available for encoding the auditory stimulus in memory (McCoy et al., 2005). Thus, listeners with poorer hearing should be less likely to update their phonetic categories, because of their poorer encoding of the stimulus sound.

Reliance on the lexicon has been shown to increase when attentional resources are taken away, for example, due to a secondary attention-demanding task (Mattys & Wiget, 2011; Mattys et al., 2013, Mattys & Scharenborg, 2014). Since lexically guided perceptual learning crucially depends on lexical knowledge, different aspects of attention or executive functioning, in particular selective attention and attention-

switching control, were being investigated in the present study as potential predictors for lexically guided perceptual learning. Attention-switching control is often measured with the Trail-Making Test, in which participants draw lines to connect numbers in a numerical sequence (e.g., 1-2-3), and in a sequence in which they constantly have to alternate between numbers and letters (e.g., 1-A-2-B-3-C). The ability to switch between these two dimensions (i.e., between the number and letter sequence) has been hypothesized to reflect a variety of cognitive processes, including the ability to remember details and to maintain two trains of thoughts simultaneously (Salthouse, 2011).

In a study on individual differences in how lexical neighborhood density affects non-word identification, Janse and Newman (2013) found that listeners with poorer attention-switching control showed larger neighborhood density effects on non-word identification. That is, non-words were easier to identify when the non-word resembled many words in the listener's language. Janse and Newman argued that listeners with poorer attention-switching control benefit more from lexical neighborhood support because they may find it particularly difficult to focus attention on unfamiliar non-word items. They concluded that listeners with poorer attention-switching control rely more strongly on top-down lexical knowledge for non-word processing than those with better attention-switching control. Consequently, in the case of lexically guided perceptual learning, where an increased reliance on lexical knowledge yields an increase in the learning effect (Scharenborg & Janse, 2013), we expected listeners with poorer attention-switching control to show more lexically guided perceptual learning than listeners with better attention-switching control. Listeners with better attention-switching control may be better able to keep both the lexical representation and the imperfectly matching acoustic form active in memory, thus to keep active the recognized word as well as keep track of a slightly imperfect match. Relationships between attentional and memory skills have been discussed in large-scale studies on cognitive control (e.g., McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010; Salthouse, Atkinson, & Berish, 2003; Miyake et al., 2000).

Selective attention also has been suggested to play a role in perceptual learning. Nosofsky and collaborators (1986, 1994), for example, suggested that listeners focus on the relevant information in the stimulus and learn to ignore irrelevant dimensions in rule-based learning of novel categories. Indeed, Janse and Adank (2012) found that listeners with better selective attention abilities showed stronger perceptual learning when presented with an artificial Dutch accent in which long vowels were systematically replaced by short vowels and vice versa. They argued that a better ability to pay attention to specific accent details helped to grasp the systematicity of the accent in their study. Although in the

present study, participants did not need to filter relevant from irrelevant stimulus characteristics, it is possible that listeners with better selective attention abilities may be drawn more strongly to the ambiguity of the target sound. Pitt and Szostak (2012) found that lexical acceptability was reduced when the attention of listeners was explicitly drawn to the signal through instructions, compared to a condition where listeners focused more on word comprehension. Pitt and Szostak gradually changed the fricative /s/ to /ʃ/ in target words, i.e., they gradually changed words containing /s/ (such as *impressive*) into their /ʃ/-containing non-word counterparts, and presented these for lexical decision. When attention was drawn away from the signal, fewer items containing ambiguous sounds were regarded as words. We therefore expected that listeners with better selective attention abilities would accept fewer of the ambiguous words as words during exposure, and consequently be less susceptible to lexically guided perceptual learning since using lexical information is critical for the retuning of phonemic categories.

Under the assumption that attention influences all modalities, including visual and auditory modalities, visual tasks of attention-switching control and selective attention rather than auditory versions were being used in the present study. Moreover, attentional abilities and hearing sensitivity were entered simultaneously into the statistical models, and in order to ensure that effects can be fully attributed to individual differences in attention-switching control or selective attention and not to individual differences in hearing sensitivity, these factors should not be correlated. We thus investigated the role of domain-general attentional abilities on lexically guided perceptual learning. Attention-switching control was investigated using the Trail-Making Test, while selective attention was investigated using a visual Flanker test.

As in previous lexically guided perceptual learning experiments, our experiment consisted of two parts (following Norris et al., 2003; Scharenborg & Janse, 2013). First, in the exposure phase, Dutch listeners were exposed to an ambiguous [f/s] sound in Dutch words ending on either /f/ or /s/ in a self-paced lexical decision task, e.g., *paradij[f/s]* (*paradijs*, “paradise”) or *witlo[f/s]* (*witlof*, “chicory”). Thus, one exposure group was exposed to words in which all word-final /f/s were replaced by an ambiguous [f/s] sound (henceforth referred to as exposure group ambF), and the other exposure group was exposed to words in which all word-final /s/s were replaced by an ambiguous [f/s] sound (henceforth referred to as exposure group ambS). In a subsequent self-paced phonetic categorization task (the test phase), listeners were confronted with a range of ambiguous sounds from an [f]-[s]-continuum appearing as the final phonemes of Dutch words, and were asked to decide whether the sound was /f/ or /s/. Both the /s/- and the /f/-reading of the ambiguous test items yielded an existing Dutch word (e.g., *brie[f/s]* could be *brief* “letter” or

bries “breeze”). The research question was whether individual differences in hearing sensitivity and attentional abilities could predict the degree of lexically guided perceptual learning in older adults.

Hearing and cognitive background tests

Individual performance on the phonetic categorization task was compared with scores on the hearing and attention tests. The perceptual learning experiment and the background tests were administered in one session with a break in the middle. Below is an overview of the experiment and background tests used in this study with the approximate duration of each task in brackets:

1. Trail-Making Test for attention-switching control (3 min)
2. Flanker task for selective attention (5 min)
3. Main experiment: Lexical decision (15 min)
4. Main experiment: Phonetic categorization (7 min)
5. Pure-tone audiogram for hearing sensitivity (10 min)

Cognitive measures

Trail-Making Test The Trail-Making Test is a paper-and-pencil task that measures executive function, or, more specifically, attention-switching control (TMT; Reitan, 1958). It consists of two parts. In part A, participants have to connect 25 digits in ascending order (i.e., 1-2-3). In part B, participants have to connect 25 digits and letters alternating between the two dimensions (i.e., 1-A-2-B-3-C). The mean time to complete part A was 43.8 s (SD = 15.4) and part B was 78.7 s (SD = 27.5). The measure used in our analysis was the ratio score of Part A and Part B (TMT-B/TMT-A). The ratio score, rather than the difference score between Parts A and B, was used to take into account general slowing (Salthouse, 2011; Verhaeghen & De Meersman, 1998). The mean ratio score on the TMT was 1.87 (SD = .59): the higher the ratio, the poorer a participant's attention-switching control. Table 4 in Appendix 1 provides separate descriptive information for the two exposure groups. Mean TMT did not differ between the two exposure groups ($t(64) < 1, p > .1$).

Flanker task The classic Flanker task (Eriksen & Eriksen, 1974) was used as a measure of selective attention. In the computerized version of this task as used in our experiment, participants have to indicate (by clicking either the ‘z’ or the ‘/’ key on the keyboard) which direction the middle symbol (a leftward or rightward pointing arrowhead) in a row of five symbols points in three types of

contexts, congruent (>>>> or <<<<<), incongruent (<<<<< or >>>>>), and neutral (==>== or ==<==). The participants were instructed to maximize speed and accuracy. Each trial started with a beep (a 400 Hz pure tone) and a fixation cross that remained on the screen for 250 ms. Following this fixation cross, the symbol string was presented for 1,500 ms. After these time, the string was removed and participants could no longer respond. The inter-trial time was 1,000 ms. Six different stimuli were each presented 12 times in the test part (the order of trial presentation was randomized for each participant) to make 72 trials. Before the test started, six practice trials were presented. The mean accuracy of the responses pooled over all participants was 92.7 % correct (SD = 26.0). Accuracy was lowest and most variable in the incongruent condition (87.3 % correct, SD = 33.3), while accuracy was best and least variable in the congruent condition (95.8 % correct, SD = 20.0); accuracy for the neutral condition was close to that of the congruent condition (95.0 % correct, SD = 21.8). Mean response times for the three conditions (calculated from visual presentation onset and collapsed only over correct responses) were: congruent condition: 572.5 ms (SD = 259.3), incongruent condition: 654.8 ms (SD = 296.0), neutral condition: 554.2 ms (SD = 223.2). The individual Flanker cost for each participant was then determined by computing the Flanker interference effect, i.e., each participant's mean logRT (of the RT in ms) in the incongruent condition was divided by that individual's mean logRT in the neutral condition. The mean logRT calculated over all participants for the incongruent condition was 6.35 (SD = .25) and 6.19 (SD = .20) for the neutral condition. The mean Flanker cost was 1.03 (SD = .04): the higher a participant's Flanker cost, the poorer was their selective attention. Table 4 in Appendix 2 provides descriptive information for the two exposure groups, separately. Mean Flanker cost did not differ between the two exposure groups ($t(60) < 1, p > .07$).

Hearing sensitivity

Hearing sensitivity was assessed with a portable Maico ST 25 screening audiometer (air conduction thresholds only, for octave frequencies from 250 Hz through 8 kHz) in a sound-attenuated booth. Given the high-frequency hearing loss associated with aging, a pure-tone average threshold was computed as the average over participants' thresholds at 1, 2, and 4 kHz. None of the participants wore hearing aids in daily life. Mean pure-tone average (PTA, in their better ear) was 17.9 dB¹ HL (SD = 11.0): the higher the participants' PTA,

¹ In The Netherlands, people are entitled to a partial refund of the costs of a hearing aid from their health insurance if this average threshold exceeds 35 dB HL in their poorer ear. So, hearing loss of the participants, if present, was mostly mild to moderate.

the poorer was their hearing sensitivity. Figure 4 in Appendix 2 shows mean pure-tone thresholds at 250–8 kHz, with the range for each threshold indicated with black dotted lines (ambF) and grey dashed lines (ambS) for the two exposure groups for the left and right ear separately, while Table 4 in Appendix 2 provides general descriptive information for the two exposure groups. Mean hearing sensitivity did not differ between the two exposure groups ($t(64) < 1, p > .1$).

Intercorrelations between participant characteristics

None of the hearing and attention background predictors were significantly correlated with each other. The highest correlation was found between attention-switching control and selective attention: Pearson's $r = .19$ ($p > .1$). The correlation between hearing loss and attention switching control was $.14$ ($p > .2$) and between hearing loss and selective attention $.06$ ($p > .6$).

Method

Participants

Sixty-six native Dutch speakers were drawn from the Max Planck Institute for Psycholinguistics subject pool and were paid for their participation. All listeners were aged 60+ years and were from the Nijmegen area (22 males; mean age: 66.7; $SD = 5.2$, age range: 60–80 years; there was no age cut-off) and had no history of severe language, speech, or hearing problems.

Materials

For the exposure phase, 20 Dutch words with final /f/ (e.g., *wiflof*, “chicory”) and 20 Dutch words with final /s/ (e.g., *paradijs*, “paradise”, see Appendix 1) were selected as target words (no further /f/ and /s/ occurred in the words); there were no minimal pairs. Syllable structure, stress patterns, and word frequency were matched as far as possible. The target items used during exposure were identical to those used in Norris et al. (2003). In addition, 60 words with no /f/ or /s/ and 100 non-words were selected as fillers (following the same syllable length distribution as for the target words): the additional words were added to disguise which items were the critical items, thereby ensuring that listeners' attention was not drawn to the ambiguous sounds, and the non-words were added to balance lexical status in the lexical decision experiment. The non-words were created such that they did

not closely resemble existing Dutch words but they did follow Dutch phonotactic rules. The non-words tended to become non-words (i.e., were no longer consistent with any real Dutch words) before their final phonemes. (Note that the filler words and non-words differed somewhat from those used in Norris et al., 2003.) Non-words were thus clear non-words, they did not only differ subtly from real words (e.g., by a single phoneme).

All words were produced in isolation by a female native Dutch speaker and digitally recorded in a sound-attenuated booth at 44 kHz. The speaker also recorded four minimal word pairs for the test phase: *brief-bries* (“letter”-“breeze”), *graf-gras* (“grave”-“grass”), *leef-lees* (“live”-“read”), and *lof-los* (“praise”-“loose”). Subsequently, ambiguous versions of the 40 critical exposure words ending in [f] and [s] as well as of the words used in the phonetic categorization task were made as follows. The critical items of the exposure stimuli contained one of eight different vowels preceding the final [f] or [s]. The selection of ambiguous sounds was done separately for each vowel. Therefore, eight minimal pairs (including the four minimal pairs used in the test phase), one for each vowel, were used to create the ambiguous sounds, and used in a pretest.

To create the ambiguous [f/s], for each of the eight minimal pairs, the final fricative was excised and zero-padded with 25 ms of silence using Praat (Boersma & Weenink, 2005), and subsequently morphed with its counterpart from the minimal pair to create an equally spaced 11-step continuum using STRAIGHT (Kawahara, Masuda-Katsuse, & Cheveigne, 1999) in Matlab. For the pretest stimuli, the ambiguous fricatives were then concatenated as final sounds to both items in the minimal pair; i.e., all 11 ambiguous [f/s] sounds were concatenated with *brie-* from source word *brief* and with *brie-* from source word *bries*.

During the pretest, six [f]-[s]-continuum steps for each of the eight minimal pairs were each presented six times binaurally over Sennheiser HD 280-13 headphones. Each stimulus was presented 500 ms after trial onset. Ten adults aged over 60 years, who did not participate in the main experiment, took part in the pretest. All participants were tested individually in a sound-treated booth. The task for the participants was to indicate by button press as quickly and as accurately as possible whether they heard the /f/-final reading or the /s/-final reading of the word. To help the listeners, the /f/-final word interpretation was always printed on the left side of the screen and the /s/-final word interpretation on the right side of the screen.

The total proportions of /s/-responses to each of the tested morphs (averaged over all eight minimal pairs) were

calculated, and the most ambiguous morph was determined, which was step 4 on the continuum. Finally, the eight-step 4 morphs were concatenated as final sounds to their vowel-consistent /f/- and /s/-final items, in the same manner as was done to create the stimuli for the pretest. This resulted in 40 stimulus pairs consisting of the same word ending in either a natural [f] or [s] or the selected ambiguous [f/s] sound. These stimuli were then used in the lexical decision task. The stimuli used for the phonetic categorization task consisted of five versions of the four minimal pairs listed above, which were created by concatenating five versions of [f/s] (i.e., steps 1, 3, 4, 5, 7) as final sound to both items of each minimal pair. Figure 1 shows the spectra of the fricative noises of the original unambiguous /f/ taken from the original *brief* source word, the ambiguous sound (morph step 4, middle panel), and of the original unambiguous /s/ taken from the original *bries* source word (bottom panel). The fricatives were taken from the word stimuli scaled to a mean intensity of 75 dB SPL.

Procedure

Two experimental-word lists were created in which the test items appeared in a pseudo-randomized running order, one for each experimental condition. The restrictions were that no critical item (i.e., no word ending in [f/s]) was allowed to appear in the first six words, and no two critical items were allowed to appear within a range of four words. Each list consisted of 200 words, i.e., 100 non-words, 60 filler words, 20 words ending in a clear [f] or [s], and 20 critical items, i.e., the /f/-final or /s/-final words ending in [f/s]. The difference between the two word lists was that one list contained only natural /f/-final words and /s/-final words ending in [f/s], the other list contained the natural /s/-final words and the /f/-words ending in [f/s]. During the exposure phase, participants were split into two groups: one group listened to the experimental-word list containing the ambiguous /f/-final words during exposure and the other group listened to the experimental-word list containing the ambiguous /s/-final words during exposure.

Participants were tested individually in a sound-treated booth. The stimuli were presented binaurally over Sennheiser HD 280-13 headphones at a fixed mean intensity level of 75 dB SPL (same intensity level for all participants). Participants were asked to press a button as quickly 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.

Next, participants were tested on the phonetic categorization task on the four minimal pairs. They were

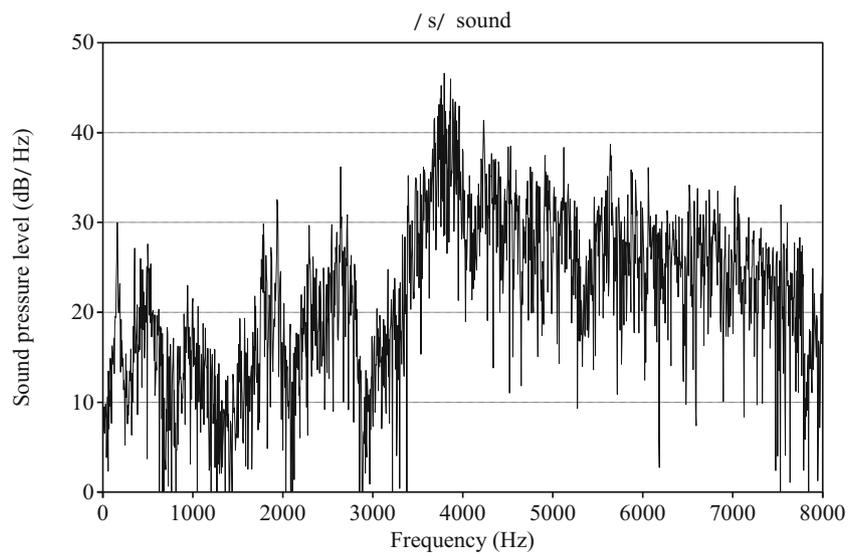
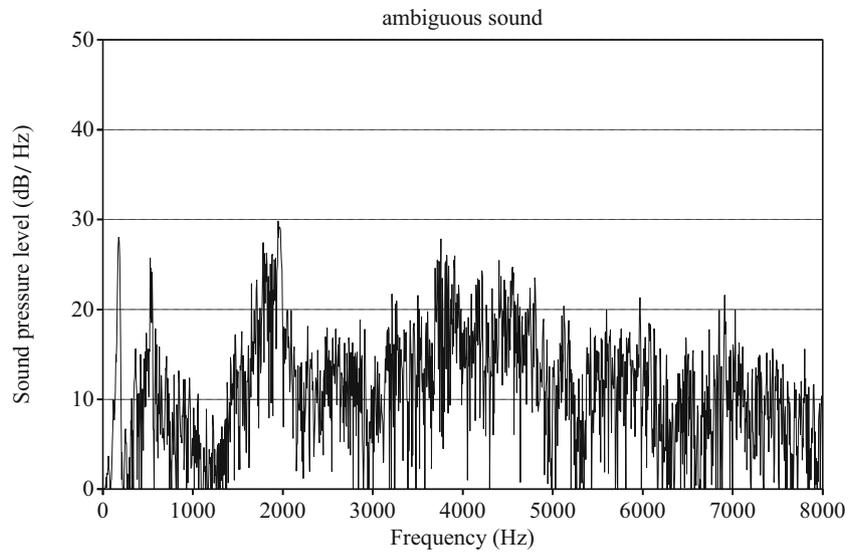
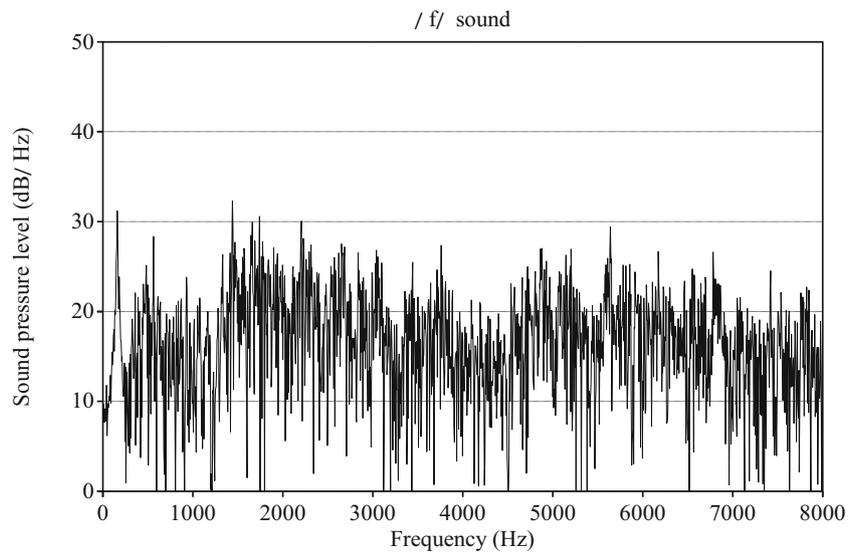
Fig. 1 Spectra of the fricative noises of /f/ from the original *brief* source word (upper panel), of the ambiguous sound (morph step 4, middle panel), and of the /s/ from the original *bries* source word (bottom panel)

asked to decide as quickly and accurately as possible, by button press, whether they heard the /f/-final reading or the /s/-final reading of the word. The five ambiguous items of each word in each minimal pair were each presented once per block (so each participant heard both source words; i.e., 40 items/block), and were newly randomized for each of a total of four blocks (160 items in total). To aid the participants, the /f/-interpretation of the stimulus was always shown on the bottom left of the computer screen, and the /s/-interpretation of the stimulus always on the bottom right.

Statistical analyses

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 for binomial data in R (R development core team, 2011). By-subject and by-minimal-pair random intercepts and random slopes were added to the models to create a maximal random effects structure (e.g., Barr, Levy, Scheepers, & Tily, 2013) and to allow for differences in size of the effect of the fixed variables under investigation between participants and stimuli (Cunnings, 2012). Note that including random slopes makes the analysis conservative by reducing the Type I error rate. The choice of fixed and random factors varied with the analysis. They are therefore listed separately for each analysis. The parameters of the generalized linear models are set using maximum likelihood estimation. We used dummy coding.

In each analysis, a best-fitting model was built using the fixed and random variables. Each analysis started by building a model containing all predictors and all possible interactions between the predictors in the fixed part of the model. Subsequently, the data was analyzed by means of a backward stepwise selection procedure, in which first interactions and then predictors that proved not significant at the 5 % level were removed one-by-one from the model, always removing the least significant interaction or predictor first. Each change in the fixed-effect structure was evaluated in terms of model fit (i.e., whether the new model explains more or less of the variance in the data compared to the previous model) by means of a likelihood ratio test with the *anova()* function in R. After the most parsimonious model which only contained significant predictors and interactions in the fixed part was determined, the best maximal random slope structure was identified. Random



intercepts and random slopes were added and tested through model comparisons, first the by-subject slope followed by the by-minimal-pair slope, if applicable. Changes in the random-slope structure were evaluated by means of the AIC (Akaike Information Criterion). The model with the lower AIC value and, therefore, better model fit was retained. The final, best-fitting model thus only contained significant predictors and interactions. In the following analyses, only these statistically significant effects are reported. In addition, we report the absolute estimated values of the different β s, which represent the relevance (effect size) of the different predictors for the estimation of the logit p (Chatterjee, Hadi, & Price, 2000) with an explanation of the effect found.

Perceptual learning results

Lexical decision

The percentage of “yes it is a word” responses to the filler words was on average 96.5 % for the listeners exposed to the ambiguous sound in the /f/-final words and 95.2 % for the listeners exposed to the ambiguous sound in the /s/-final words. The percentage of “yes it is a word” responses to the non-word filler items was 3.8 % and 4.7 % for the listeners exposed to the ambiguous sound in the /f/-final words and the /s/-final words, respectively. Table 1 shows the performance on the lexical decision task as mean percentages of “yes” responses to the natural and ambiguous versions of the /f/- and /s/-final target words.

The target items were subsequently analyzed with the response to the target item being the dependent variable (“yes it is a word”: coded as 1; “not a word”: coded as 0). Fixed variables were type of stimulus (word ending in the natural vs. ambiguous fricative) and exposure group (exposed to the ambiguous sound only in /f/-final words during the lexical decision task (ambF) or only in /s/-final words (ambS)); subject and item were random variables.

Table 1 Performance on the lexical decision task as mean percentages of “yes” responses to the natural and ambiguous versions of the /f/- and /s/-final target words

Mean % yes	Natural fricatives		Ambiguous fricatives	
	/s/-final	/f/-final	/s/-final	/f/-final
Target words	95.6	98.0	91.2	89.2

The difference in “yes” responses to the natural stimuli and the ambiguous stimuli was significant ($\beta = 1.6034$, $SE = .4916$, $p < .001$), i.e., there were significantly more “yes” responses to the natural stimuli than to the ambiguous stimuli. Nevertheless, for both exposure groups, listeners accepted the vast majority of the stimuli ending in the ambiguous sound as words (see also Table 1), i.e., listeners who were exposed to the ambiguous [f/s] in the normally /s/-final words tended to interpret the ambiguous sound as /s/, whereas listeners who were exposed to the ambiguous sound in the context of normally /f/-final words interpreted [f/s] as /f/. The maximal random slope structure of the final model included random effects for subject and minimal pair and a participant random slope for minimal pair. Inclusion of the random slope suggests that the difference in ‘yes’ responses to words ending in a natural versus an ambiguous fricative differs between participants.

Phonetic Categorization

Figure 2 shows the proportion of /s/ and /f/ responses for the five ambiguous stimuli in the phonetic categorization task averaged over the four test blocks. The responses for the listeners who were exposed to [f/s] in the normally /s/-final words are indicated with ‘S’. The responses for the listeners who were exposed to [f/s] only in the normally /f/-final words are indicated with

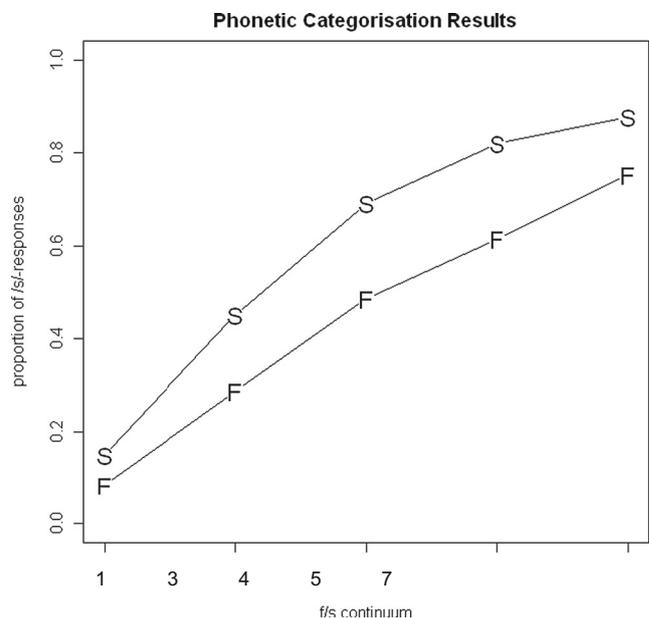


Fig. 2 The total proportion of /s/ responses for the two exposure conditions: ‘S’ indicates the group of listeners who learned to map [f/s] onto [s]; ‘F’ indicates the group of listeners who learned to map [f/s] onto [f] for the five ambiguous test stimuli

'F'. The responses to the five ambiguous stimuli were analyzed statistically (the dependent variable was whether the response is /f/, coded as 0, or /s/, coded as 1). The fixed variable was the exposure group (ambF on the intercept). Stimulus step (step 1 through 7, see Fig. 2; note that the stimulus step is a continuous variable, so the stimulus steps can be seen as lying on a straight line from step 1 to step 7; step 4 is on the intercept) was used as the control variable. Subject and minimal pair were the random factors. Table 2 displays the parameter estimates in the best-fitting model of performance.

As can be seen in Fig. 2, listeners who were exposed to [f/s] in the normally /s/-final words (ambS) were biased to label the sounds on the continuum as /s/, while those listeners who were exposed to [f/s] in the normally /f/-final words (ambF) were less likely to do so (see also the significant effect of Exposure group in Table 2). The difference between the curves of the mean proportion of /s/ responses for the two exposure groups in Fig. 2 is taken to be the perceptual learning effect. Thus, older listeners were generally capable of perceptual learning of ambiguous sounds on the /f-/s/ continuum. Significantly more /s/ responses were given to higher stimulus steps on the continuum, which is to be expected as 'higher' means more /s/-like ambiguous sounds (Table 2: Stimulus step), and this was even more so the case for the listeners exposed to the ambiguous sound in the /s/-final words (Exposure group × Stimulus step). The inclusion of a participant random slope for stimulus step indicates that the degree of increase in /s/ responses differed across participants. Moreover, the lack of a minimal-pair random slope for exposure condition indicates that the effect of exposure condition was not driven by some minimal pairs, but generalizes across minimal pairs.

Individual differences in the perceptual learning effect

In order to investigate the effect of hearing sensitivity and attentional abilities on lexically guided perceptual

learning, variation in the size of the perceptual learning effect as well as variation in the predictor variables was needed. Since the end points of the continuum showed only a small difference between the two exposure groups (see also Fig. 2), we only looked at the most ambiguous stimulus steps (i.e., steps 3, 4, and 5 in Fig. 2). A new category "learning-consistent" was created. The trials in the phonetic categorization task were relabelled such that an /f/ response by listeners exposed to the ambiguous sound in /f/-final words and an /s/ response by listeners exposed to the ambiguous sound in the /s/-final words were labelled as '1', i.e., "learning-consistent", while trials that were "learning-inconsistent" were relabelled as '0'. We predicted the category "learning-consistent" in our analysis. Fixed variables were hearing sensitivity, selective attention, and attention-switching control (centralized around the mean). Subject and minimal pair were the random factors, while exposure group was entered as a control variable. Table 3 displays the parameter estimates in the best-fitting model of performance.

Figure 3 visualizes the relationships between learning consistency on the one hand (calculated as the percentage learning-consistent responses for each listener individually) and the hearing and attentional measures introduced earlier on the other hand, i.e., hearing sensitivity (the pure-tone average threshold (PTA) computed as the average over each participant's thresholds at 1, 2, and 4 kHz in their better ear), selective attention (each individual's Flanker interference effect, i.e., each participant's mean logRT in the incongruent condition divided by that individual's mean logRT in the neutral condition), and attention-switching control (the ratio score of Part A and Part B of the Trail-Making Test for each participant individually). For ease of interpretation, the scores for hearing sensitivity, selective attention, and attention-switching control are converted into z-scores. Note again that for all individual measures, higher values on the x-axis mean poorer performance on the particular background test.

The final model showed that only attention-switching control had a significant effect on learning consistency: more

Table 2 Fixed effect estimates for the best-fitting model of performance in the phonetic categorization task (n = 10560). The group of listeners exposed to the ambiguous words in /f/-final words are on the intercept

Fixed effect	β	Standard error	$p <$
<i>Intercept</i>	-.4237	.4662	n.s.
Exposure group	1.3869	.3700	.001
Stimulus step	.8407	.0982	.001
Exposure group × Stimulus step	.2992	.1400	.05

Table 3 Fixed effect estimates for the best-fitting model of learning consistency performance in the individual differences analysis (n = 6336). The group of listeners exposed to the ambiguous words in /f/-final words are on the intercept

Fixed effect	β	Standard error	$p <$
<i>Intercept</i>	.2578	.4454	n.s.
Exposure group	.5585	.7645	n.s.
Attention-switching control	.5123	.2477	.05

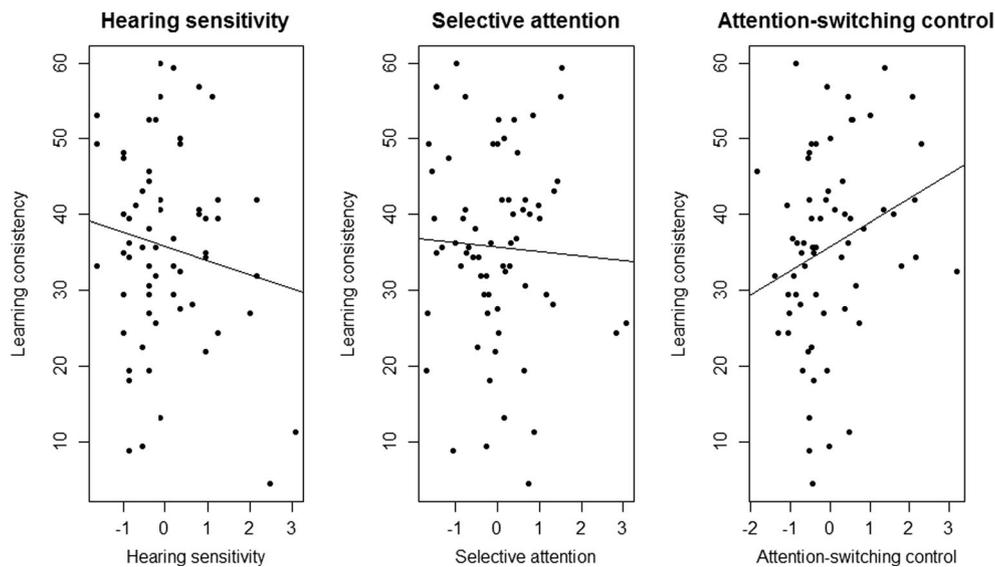


Fig. 3 Scatter plots with regression lines of percentage of learning-consistent responses as a function of the (z-scores of the) hearing and attention measures: hearing sensitivity, selective attention, and attention-switching control, for the two exposure groups together. The only significant correlation is that between percentage of learning-consistent

responses and attention-switching control (Spearman's $\rho = .29$, $p < .05$; correlation with hearing sensitivity: Spearman's $\rho = -.01$, $p > .9$; correlation with selective attention: Spearman's $\rho = .03$, $p > .8$). Note that higher values on the x-axis mean poorer performance for all three background tests

learning-consistent responses were given with poorer attention-switching control (see Table 3: Attention-switching control, and the upward-trending regression line in Fig. 3). Thus, only attention-switching control modified the strength of the perceptual learning effect.² A median split comparison of the participants into a group of participants with better attention-switching control (TMT ≤ 1.66) and a group of participants with worse attention-switching control (TMT > 1.66) confirmed that the two groups of participants differed significantly in their number of learning consistent responses ($t(64) = -2.55$, $p < .01$).

Neither hearing nor selective attention modified the lexically guided perceptual learning effect. Moreover, there was no effect of exposure condition on learning consistency (see Table 3: Exposure group); i.e., there was no difference in the number of learning-consistent responses between the two groups of listeners. The degree to which attention-switching control predicted the lexically guided perceptual learning effect differed for the four minimal pairs that were used as shown by a minimal pair random slope for attention-switching control. Since the simple effect of attention-switching control in the fixed structure remained significant, our attentional effect on learning-consistent categorization

behavior held generally over minimal pairs, and was thus robust. Moreover, the participants in the two exposure groups showed differences in the number of learning-consistent responses to the different minimal pairs as shown by a minimal-pair random slope for exposure condition. Importantly, despite the conservative statistical methodology of including random slopes that was used here, there is a general effect of attention-switching control. Although this effect is not particularly large, the result shows that individual differences in attentional control are significantly associated with the way participants process the speech signal.

The spectra in Fig. 1 show that part of the information distinguishing the two fricatives is around 4000 Hz. We therefore reran the above analysis to investigate a participant's hearing threshold in their better ear at 4000 Hz as a predictor for learning consistency. This analysis showed the same result as the above analysis: only attention-switching control modified the lexically guided perceptual learning effect. A participant's hearing threshold at 4000 Hz did not.

General discussion

This study investigated two possible variables that might be involved in lexically guided perceptual learning: hearing sensitivity and attentional abilities – more specifically, selective attention and attention-switching control. We expected lexically guided perceptual learning to be dependent on listeners' hearing abilities, as positive evidence that a sound is ambiguous is critical for category retuning. Furthermore, listeners

² We reran the analysis excluding all (three, outlier) poorer-hearing participants with a PTABest of over 35 dB HL to investigate whether there is an effect of selective attention for the people with better hearing which might have been obscured by the results of the listeners with poor hearing. The results were the same as the results reported here, i.e., only an effect of attention-switching control was found, while still no effect was found for selective attention.

with better selective attention may be drawn to the ambiguity of the target word's final sound, leading to less susceptibility to lexically guided perceptual learning. Attention-switching control may change the weighting of bottom-up acoustic information versus lexical information. Since the lexicon drives category retuning, we expected attention-switching control to modify the perceptual learning effect such that listeners with poorer attention-switching control are more susceptible to lexically guided perceptual learning effects given their increased reliance on lexical information.

We present the results of a group of older (aged 60+ years) Dutch listeners on lexically guided perceptual learning. The main experiment consisted of two parts: an exposure phase consisting of a lexical decision phase and a phonetic categorization test phase. During the exposure phase, listeners learned to interpret the ambiguous final [f/s] in Dutch words as /f/ or /s/ depending on the lexical context in which they encountered these sounds.

For lexically guided phonetic category retuning to occur, an ambiguous sound should at the same time be interpretable as the sound expected on the basis of the lexical representation and deviate from it. In our experiment, the percentages of 'yes' responses to the ambiguous items averaged over the two exposure groups was 90.2 %, which is highly similar to the percentage of 'yes' responses of older listeners to the ambiguous items on a liquid contrast (89.8 %; Scharenborg & Janse, 2013) and also to that of younger listeners on the /f/-s/ contrast (90.0 %, Norris et al., 2003). The older listeners in our study thus overwhelmingly accepted the items containing the ambiguous fricative as words, which suggests that the fricative sound is indeed generally a "good enough" representative of either /s/ or /f/, at least in the face of overwhelming lexical evidence for one specific word. Nevertheless, the sound was deviant enough for participants to provide them with the necessary evidence for phonetic category retuning. Since the formant transitions from the preceding vowel into the final fricative were not morphed and were thus congruent with the lexical evidence for the final fricative's identity during the exposure phase, it must have been the fricative noise that made the ambiguous [f/s] sound deviant from either [f] or [s] and induced lexically guided category retuning. Note that in many cases, listeners did not strictly need the final phoneme to recognize the word: many of the monosyllabic words and all of the multisyllabic words used in the present study become unique before the final phoneme.

During the phonetic categorization test phase, those listeners who were exposed to the ambiguous sound in /f/-final words more often labelled the ambiguous test items as /f/, whereas listeners exposed to the ambiguous sound in /s/-final words more often labelled the ambiguous test items as /s/. As Norris et al. (2003) found for young adult listeners, older listeners show perceptual learning of ambiguous sounds on the [f]-[s] continuum.

Unexpectedly, hearing loss in our participant sample did not seem to modify the size of lexically guided perceptual learning. As mentioned before, during the exposure phase listeners did not need the final fricative to make their lexical decision responses as most target words became unique before the final phoneme. However, there is only a need for lexical retuning if the participant hears a sound that deviates from his or her expectations regarding the sound. We assumed that listeners with hearing loss would be less able to hear the ambiguity of the sound given the fricative contrast used in this study. This is not what we found, which is unexpected if one compares the audiogram provided in Fig. 4 (Appendix 2) with the spectra of /f/ and /s/ and the ambiguous fricative given in Fig. 1. From these figures, it is clear that energy in the presented fricative sounds was generally only between 10 and 30 dB SPL, and that most of the information distinguishing the two fricatives is between 4000 and 5000 Hz. Given that the average hearing threshold at 4000 Hz was around 25 dB HL (ranging from -5 to 75 dB HL), the difference between the two fricatives at 4000 Hz must have been inaudible to some of our participants. We can only speculate that listeners may have attended to the difference in energy distribution in a lower frequency range (e.g., up to 2000 Hz). In the lower frequency range, hearing thresholds were more favorable than at 4000 Hz, such that listeners may have still been able to distinguish the fricative sounds. Alternatively, the lack of a relationship between hearing sensitivity and category retuning may indicate that the hearing sensitivity measure used here may not have been the most relevant index of a participant's sensitivity to this phoneme contrast. Other hearing-related differences between participants (e.g., auditory processing measures such as frequency resolution) may have been more relevant for fricative perception, and hence lexically guided perceptual learning. In short, our results show that even listeners with some hearing loss in the higher frequency range show a perceptual learning effect with the [f]-[s] contrast: lexically guided perceptual learning is thus not completely lost when acoustic processing is less accurate (Cutler, 2012). Finally, we should note that high-frequency hearing loss may possibly lead to broadening of multiple or even all phoneme categories. This broadening, however, is different from category retuning induced by lexically guided perceptual learning, which is sound-specific (Mitterer et al., 2013) and context-specific (Jesse & McQueen, 2011; Mitterer et al., 2013).

Several studies report that participants with better selective attention show more perceptual learning. Both for rule-based learning of novel categories (Nosofsky, 1986; Nosofsky, Gluck, Palmeri, McKinly, & Glauthier, 1994) and for learning of an artificial Dutch accent (Janse & Adank, 2012), better selective attention abilities were suggested to lead to better learning. In the context of lexically guided perceptual learning, we suggested that listeners with better selective attention may be drawn to the ambiguity of the target word's final sound, which would reduce the listener's susceptibility to lexically guided perceptual learning. We, however, found no effect of selective

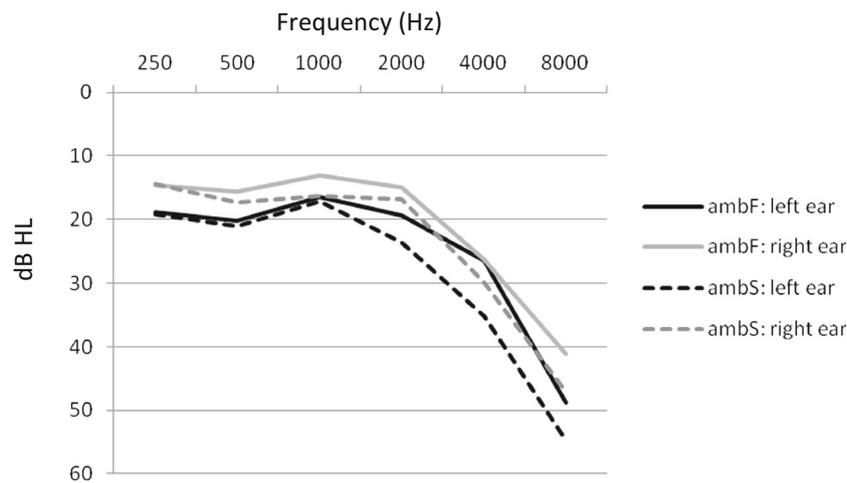


Fig. 4 Mean hearing sensitivity in terms of pure-tone hearing thresholds from 250 to 8000 kHz of the listeners exposed to the ambiguous sound in /f/-final words (ambF) and of the listeners exposed to the ambiguous sound in /s/-final words (ambS), for the two ears separately

attention on lexically guided perceptual learning. This raises the question whether attentional abilities were associated with a participant's lexical decision task performance during the exposure phase. A separate analysis investigating the role of hearing loss, selective attention, and attention-switching control on the percentage of 'yes' responses during the lexical decision task showed no role for any of these hearing and cognitive measures. Thus, contrary to our hypothesis, listeners with better selective attention abilities did not accept fewer of the ambiguous words as words and consequently they were not less susceptible to lexically guided perceptual learning than listeners with worse selective attention. This may however have been due to the relatively small variation in the percentage of 'yes' responses to the ambiguous items (61 out of the 66 participants had percentages above 80 %; mean: 90.2 %).

Participants quickly adjust their phoneme category boundaries upon exposure to a novel, ambiguous sound. In the introduction, we hypothesized that having poorer attention-switching control would lead to an increased lexically guided perceptual learning effect, due to an increased reliance on lexical information compared to having better attention-switching control. This is indeed what we found. Possibly, poorer attention-switching ability limits listeners' attention to be drawn to the ambiguous sound, particularly in the face of overwhelming evidence for one particular lexical candidate. This would in turn enhance lexically guided adjustment of the phoneme category, such that listeners with poorer attention-switching abilities would retune more than listeners with better attention-switching abilities. This link between reliance on lexical information and the strength of the lexically guided perceptual learning effect has already been made explicit by Scharenborg and Janse (2013). They found, using a lexically guided perceptual learning task of liquid consonants, that listeners who more often judged an ambiguous item as a word in the lexical decision task had a stronger lexically guided perceptual learning effect compared to listeners who judged an ambiguous item as a word less often. Thus, listeners who gave more

learning-consistent responses showed stronger perceptual learning than listeners who regarded an ambiguous item as a word less often. Note, however, that in our study, no effect of attention-switching control was found on lexical decision task performance. Mattys and colleagues (2011, 2013) have shown in a series of experiments that the relative weight of lexical versus acoustic information changes in different listening/testing conditions. They provide two possible accounts for their results (2013). First, anxiety (imposed through a carbon dioxide anxiety-inducing manipulation in Mattys et al., 2013) and divided attention (Mattys & Wiget, 2011) may distract listeners away from the acoustic detail of the signal, such that they fall back on lexical knowledge. Our results are not entirely consistent with this explanation: listeners with poorer attention-switching control rely more on lexical knowledge, but at the same time incorporate the idiosyncratic pronunciation of the sound in their category (i.e., they have more learning-consistent responses). Our results are more consistent with a second explanation offered by Mattys and colleagues which says that anxiety and cognitive load impair inhibitory control, such that listeners are less able to inhibit lexical activation, maximizing their susceptibility to lexically guided retuning. Possibly, listeners with better cognitive abilities may either more easily change how they weight bottom-up and top-down information in speech perception, or have a permanently different weighting compared to listeners with poorer attention switching. These listeners with better attention-switching control are then biased toward a more bottom-up mechanism of perception, i.e., they rely relatively more on interpretations based on signal properties rather than on top-down, knowledge-guided heuristics, or are better able to switch back to the acoustic form representation (possibly by inhibiting lexical activation) once the target word has been recognized. Phrased differently, listeners with better attention-switching control may be better able to keep both the acoustic form and the lexical representations active in memory, and to switch back to the acoustic form representation and keep track of a slightly

imperfect match. This keeping track of the slight mismatch and hence reduced reliance on lexical knowledge makes them in turn less susceptible to lexically guided perceptual learning effects, as shown here. On the other hand, listeners with poorer attention-switching control may be less able to keep in memory the imperfect match of the target word with the stored representation, making these listeners maximally susceptible to lexical guidance for category retuning. Our results also concur with those found by Janse and Newman (2013), who had participants identify monosyllabic non-words differing in the number of lexical neighbors. Greater neighborhood density generally facilitated non-word identification, but more so for participants with poorer attention-switching control. They argued that listeners with poorer attention-switching control may find it particularly difficult to focus attention on an unfamiliar non-word item, and benefit more from neighborhood support. Both the present perceptual learning effects and the neighborhood results of Janse and Newman (2013) therefore suggest that those with poorer attention-switching control rely more strongly on top-down lexical evidence in speech perception than those with better attention-switching control. Our finding of the role of attention-switching control on lexically guided perceptual learning may thus well have a bearing not only on lexically guided perceptual learning, but on the manner in which lexical knowledge mediates speech perception more generally. Our results seem to suggest that when a speaker has a subtle speech impediment that does not hamper understanding the words in the message, listeners with poorer attention-switching control might not notice and may later not remember as well as listeners with better attention-switching control that the speaker actually had a speech impediment.

The picture that arises from our results is that in order for lexically guided perceptual learning to have its maximal effect, listeners' attention should *not* be diverted to the deviant sound, suggesting that lexically guided perceptual learning is a type of "implicit learning." Implicit learning has been argued to occur with minimal demands on attention (see for an overview, Shanks, 2003), which implies that no central processing resources are needed for implicit learning. However, this claim is controversial (Shanks, 2003). The stronger reliance on lexical knowledge that was found in this study could be linked to a worse ability to inhibit lexical activation, or a worse ability to switch back to the acoustic form once there is overwhelming evidence for one particular word candidate. Note that these participants' focus on the lexical plausibility of the words (during the exposure phase) does not prevent them from updating their sound categories so as to include this new pronunciation. In other words, in as far as participants are "distracted away" from the deviant acoustic realization during the exposure phase, category retuning nevertheless occurs. However, by investigating individual differences in lexically guided perceptual learning, we allowed for the possibility that this form of "implicit" learning does relate to cognitive abilities, i.e., that individual differences may arise from differences in allocation of these resources.

Our results seem to suggest that attention is associated with, at least, this type of "implicit learning." Individual differences might actually not only play a role in perceptual learning but might be important for speech perception more generally, e.g., how individual listeners weight bottom-up information against lexical information. Moreover, lexically guided perceptual learning is a specific type of perceptual learning. Perceptual learning is thus a "container" word for many different types of learning. Goldstone (1998) provides an overview of several types of perceptual learning, including categorical perception (i.e., it is easier to distinguish between two sets of equidistant stimuli when the two stimuli are from different categories than when they are from the same category), exemplar-based learning of a stimulus (i.e., each stimulus/exemplar is stored in the brain) or of a feature of a stimulus, or differentiation of stimuli or categories. Possibly, the cognitive abilities underlying these learning processes differ, perhaps depending on how explicit or implicit the learning processes are (Unsworth & Engle, 2005).

In conclusion, this paper is the first to investigate individual differences in lexically guided perceptual learning, considering two possible driving forces for lexically guided perceptual learning: hearing sensitivity and attentional abilities. Our experiment using the /f/-/s/ contrast showed that selective attention and hearing loss did not modify perceptual learning. Lexically guided perceptual learning is thus not lost when acoustic processing is less accurate. Attention-switching control was found to be associated with perceptual learning: listeners with poorer attention-switching control have a stronger perceptual learning effect. In line with other literature on the relationship between attentional capacity and relative reliance on lexical information (Janse & Newman, 2013; Mattys & Wiget, 2011; Mattys et al., 2013), we postulate that listeners with better attention-switching control may weight bottom-up information more strongly than top-down information in speech perception compared to listeners with poorer attention-switching control, or are better able to keep in mind that there was a slight mismatch of the acoustic input with the stored representation. This reduced reliance on lexical knowledge makes listeners with better attention-switching control less susceptible to lexically guided perceptual learning effects. The lexical content of the speech signal is thus not only important for understanding the meaning of the utterance, but also for being able to quickly adapt to new speakers, unfamiliar accents, and speaker idiosyncrasies which allows listeners to recognize future words produced by the same speaker more rapidly.

Author Note This research was carried out while O.S. and A.W. were at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands. The research by O.S. was sponsored by the Max Planck International Research Network on Aging. O.S. is now supported by a Vidi-grant from the Netherlands Organisation for Scientific Research (NWO). The research by E.J. is supported by a (separate) Vidi-grant from NWO. The research by A.W. was funded by the Max Planck Society, Munich.

Germany. A.W. is now at the Eberhard Karls University of Tübingen. We thank the student-assistants of the Adaptive Listening Group of the Max Planck Institute for Psycholinguistics and Willemijn van den Berg for

assistance in preparing and running these experiments, and Marijt Witteman for recording the stimuli.

Appendixes

Appendix 1

Table 4 Descriptive results for the hearing and attention measures for the two exposure groups presented separately. ‘ambF’ refers to the listeners exposed to the ambiguous sound in /f/-final words; ‘ambS’ refers to the listeners exposed to the ambiguous sound in /s/-final words

Exposure group	Hearing sensitivity			Selective attention			Attention-switching control		
	Mean PTA _{1,2,4} kHz (dB HL)	SD	range	mean	SD	range	mean	SD	range
ambF	16.6	12.0	0 – 51.7	1.04	.05	.96 – 1.16	1.88	.62	1.09 – 3.77
ambS	19.1	10.0	0 – 41.7	1.02	.04	.95 – 1.09	1.87	.57	.79 – 3.24

Appendix 2

Results for the hearing and cognitive measures for the two exposure groups given separately. Figure 4 shows mean hearing sensitivity in terms of hearing thresholds from 250 to 8000 kHz of the listeners exposed to the ambiguous sound in /f/-final words (ambF, solid lines) and for the listeners exposed to the ambiguous sound in /s/-final words (ambS, dashed lines). The left ear is plotted in black, the right ear is plotted in grey. Descriptive results for the hearing and cognitive measures for the two exposure groups presented separately are provided in Table 4.

Table 5 Overview of the 20 /f/-final and 20 /s/-final words used in the lexical decision task

/f/-final words			
1-syllable	2-syllables	3-syllables	4-syllables
braaf	aanhef	achterneef	choreograaf
druif	karaf	biograaf	indicatief
lijf	octaaf	middenrif	kwalitatief
proef	olijf	onderlijf	landbouwbedrijf
rif	witlof	ongeloof	locomotief
/s/-final words			
1-syllable	2-syllables	3-syllables	4-syllables
baas	hakmes	geitenkaas	anekdotisch
krijs	karkas	grandioos	champagneglas
muis	naaldbos	hagedis	democratisch
nis	radijs	paradijs	ingenieus
roes	relaas	pimpelmees	problematisch

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.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.
- 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.
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*(3), 369–382.
- Cutler, A. (2012). Native listening: The flexibility dimension. *Dutch Journal of Applied Linguistics, 1*(2), 169–187.
- Cutler, A., McQueen, J. M., Butterfield, S., & Norris, D. (2008). Prelexically-driven perceptual retuning of phoneme boundaries. *Proceedings of Interspeech, 2056–2056*.
- Davis, M. H., Johnsruide, I. S., Hervais-Adelman, A., Taylor, K., McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134*(2), 222–241.
- Eisner, F., & McQueen, J. M. (2005). Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America, 119*, 1950–1953.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effect of noise letters on upon the target identification in a non-search task. *Perception and Psychophysics, 16*, 143–149.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 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. (2010a). Recognition of accented English in quiet by younger

- normal-hearing listeners and older listeners with normal-hearing and hearing loss. *Journal of the Acoustical Society of America*, 128(1), 444–455.
- Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010b). Recognition of accented English in quiet and noise by younger and older listeners. *Journal of the Acoustical Society of America*, 128(5), 3152–3160.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception & Performance*, 34(2), 460–474. doi:10.1037/0096-1523.34.2.460.
- Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *Quarterly Journal of Experimental Psychology*, 65(8), 1563–1585. doi:10.1080/17470218.2012.658822
- Janse, E., & Newman, R. S. (2013). Identifying non-words: Effects of lexical neighborhoods, phonotactic probability, and listener characteristics. *Language and Speech*, 56, 421–441.
- Jesse, A., & McQueen, J. M. (2011). Positional effects in the lexical retuning of speech perception. *Psychonomic Bulletin & Review*, 18, 943–950. doi:10.3758/s13423-011-0129-2
- 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.
- Maniwa, K., Jongman, A., & Wade, T. W. (2008). Perception of clearfricatives by normal-hearing and simulated hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 123, 1114–1125. doi:10.1121/1.2821966
- Mattys, S., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, 59, 203–243.
- Mattys, S. L., & Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors. *Psychology and Aging*, 29(1), 150–162. doi:10.1037/a0035387
- Mattys, S. L., Seymour, F., Attwood, A. S., & Munafò, M. R. (2013). Effects of acute anxiety induction on speech perception: Are anxious listeners distracted listeners? *Psychological Science*, 24(8), 1606–8. doi:10.1177/0956797612474323
- Mattys, S., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65, 145–160.
- McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24, 222–243.
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology*, 58A, 22–33.
- 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. (2012). Lexical retuning of children's speech perception: Evidence for knowledge about words' component sounds. *Language Learning and Development*, 8, 317–339. doi:10.1080/15475441.2011.641887
- Mitterer, H., Scharenborg, O., & McQueen, J. M. (2013). Phonological abstraction without phonemes in speech perception. *Cognition*.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47(2), 204–238.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 1, 39–57.
- Nosofsky, R. M., Gluck, M., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, 22, 352–369.
- 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.
- Pitt, M. A., & Szostak, C. M. (2012). A lexically biased attentional set compensates for variable speech quality caused by pronunciation variation. *Language and Cognitive Processes*, 27(7/8), 1225–1239.
- R development core team (2011). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Reitan, R. M. (1958). Validity of the Trail Making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Rietveld, A. C. M., & van Heuven, V. J. (1997). *Algemene Fonetiek*. Bussum: Dick Coutinho.
- Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence*, 39, 222–232.
- Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, 132, 566–594.
- Samuel, A. G., & Kraljic, T. (2009). Perceptual learning in speech perception. *Attention, Perception & Psychophysics*, 71, 1207–1218.
- Scharenborg, O., & Janse, E. (2013). Comparing lexically-guided perceptual learning in younger and older listeners. *Attention, Perception, and Psychophysics*. doi:10.3758/s13414-013-0422-4
- Shanks, D. R. (2003). Attention and awareness in “implicit” sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 11–42). Amsterdam: Benjamins.
- Sher, A. E., & Owens, E. (1974). Consonant confusions associated with hearing loss above 2000 Hz. *Journal of Speech and Hearing Research*, 17, 669–681.
- Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, 33(2), 213–220.
- Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging*, 13, 120–126.
- Zeng, F. G., & Tumer, C. W. (1990). Recognition of voiceless fricatives by normal and hearing-impaired subjects. *Journal of Speech and Hearing Research*, 33, 440–449.