

# Age and hearing loss and the use of acoustic cues in fricative categorization

Odette Scharenborg<sup>a),b)</sup>

Centre for Language Studies, Radboud University Nijmegen, Erasmusplein 1, 6525 HT, Nijmegen, The Netherlands

Andrea Weber<sup>b)</sup>

Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands

Esther Janse<sup>b)</sup>

Centre for Language Studies, Radboud University Nijmegen, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands

(Received 9 September 2014; revised 21 July 2015; accepted 21 July 2015; published online 10 September 2015)

This study examined the use of fricative noise information and coarticulatory cues for categorization of word-final fricatives [s] and [f] by younger and older Dutch listeners alike. Particularly, the effect of information loss in the higher frequencies on the use of these two cues for fricative categorization was investigated. If information in the higher frequencies is less strongly available, fricative identification may be impaired or listeners may learn to focus more on coarticulatory information. The present study investigates this second possibility. Phonetic categorization results showed that both younger and older Dutch listeners use the primary cue fricative noise and the secondary cue coarticulatory information to distinguish word-final [f] from [s]. Individual hearing sensitivity in the older listeners modified the use of fricative noise information, but did not modify the use of coarticulatory information. When high-frequency information was filtered out from the speech signal, fricative noise could no longer be used by the younger and older adults. Crucially, they also did not learn to rely more on coarticulatory information as a compensatory cue for fricative categorization. This suggests that listeners do not readily show compensatory use of this secondary cue to fricative identity when fricative categorization becomes difficult. © 2015 Acoustical Society of America.

[<http://dx.doi.org/10.1121/1.4927728>]

[ICB]

Pages: 1408–1417

## I. INTRODUCTION

Age-related decline in hearing particularly affects the higher frequencies. Although the definition of “high-frequency energy” is not set (different authors use different criteria, e.g., frequencies higher than 2 kHz or even 8 kHz; Monson *et al.*, 2014), the importance of high-frequency energy for the perception of speech is clear (for an overview, see Monson *et al.*, 2014). Removal of high-frequency energy, e.g., all energy above 3.5 kHz (Apoux and Bacon, 2004), affects both the perceived quality (Moore and Tan, 2003) and intelligibility (Lippmann, 1996; Moore *et al.*, 2010; Stelmachowicz *et al.*, 2001) of speech. One would therefore expect that age-related hearing loss interferes with the recognition of consonants that have their distinctive information in higher frequencies, i.e., fricatives. Fricatives have been found to frequently induce recognition errors both in quiet and in noise (e.g., Miller and Nicely, 1955; Gelfand *et al.*, 1986) and for normal-hearing listeners and listeners with (simulated) sensorineural hearing loss (e.g.,

Maniwa *et al.*, 2008; Sher and Owens, 1974; Zeng and Turner, 1990). Moreover, the recognition of fricatives deteriorates with age even in normal-hearing adults (Gelfand *et al.*, 1986). In this study, we use the Dutch consonants [f] and [s] (note that Dutch, unlike English, does not have [θ] or [ʃ]—although the latter does appear in loan words and in morphological complex words (/vis/ + /jə/ (fish + diminutive) becomes /viʃə/ (little fish)). The noise spectra of the two fricatives differ: while [s] has more energy in the higher frequencies, concentrating around 5500 Hz, the energy for [f] is distributed more uniformly throughout the spectrum (Jongman *et al.*, 2000; Rietveld and van Heuven, 1997).

In Dutch, formant transitions from the preceding vowel into a subsequent [f] or [s] differ in the slope of F3 around 2200 Hz (a rise into a following [s] and flat into a following [f]; Rietveld and van Heuven, 1997). This formant transition information is thus present at a lower frequency than the information of the spectral peak (around 5500 Hz) for [s]. If information in the higher frequencies is indeed less strongly available to older listeners with hearing loss, fricative identification may be impaired, or listeners may have learned to use other cues in the speech signal for fricative identification. The present study investigates this second possibility by looking at the use of other cues for word-final fricative

<sup>a)</sup>Electronic mail: O.Scharenborg@let.ru.nl

<sup>b)</sup>Also at: Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands.

categorization in Dutch. Fricative identification is not only dependent on the fricative noise and the formant transitions leading into the fricative (e.g., Whalen, 1981), but also influenced by the preceding vowel (Whalen, 1981) and the amplitude of the fricative noise relative to the amplitude of the vowel (Hedrick and Ohde, 1993). In this study, we specifically look at the role of coarticulatory cues in the vowel preceding word-final fricatives on fricative categorization. A potential fallback to reliance on coarticulatory cues for fricative categorization might be expected as it is a strategy employed in first language acquisition. In line with developmental results for English (Nittrouer, 1992, 2002), young Dutch children also use formant transition information for fricative identification (Gerrits, 2001), while the use of formant transition information is strongly reduced in favor of spectral noise information in older children and adults (Gerrits, 2001). Thus, older adults possibly return to the use of coarticulatory information in the preceding vowel such as formant transition information for distinguishing phonetic categories in their native language in the face of hearing loss.

In a series of phoneme monitoring experiments with pseudowords in which formant transitions for [s] and [f] were either correct or misleading, Wagner *et al.* (2006) found that misleading formant transitions did not affect how often target fricatives were detected by young, normal-hearing Dutch listeners. Moreover, the misleading formant transitions led only to a very small increase in phoneme monitoring times, prompting the authors to conclude that Dutch listeners were not affected by misleading formant transitions for fricative identification. Similarly, a gating study by Wagner (2013) showed that Dutch listeners need information from the start of the fricative noise for correct fricative identification. Note that listeners from language backgrounds with richer fricative inventories appear to benefit more from an integration of coarticulatory information from the vowel with the onset of the fricative noise for fricative identification. Together these results suggest that young, normal-hearing Dutch listeners rarely rely on formant transitions for fricative identification. However, it is possible that in the face of deteriorating hearing sensitivity, Dutch listeners start to use coarticulatory information from the preceding vowel, such as formant transitions, for fricative identification. In fact, when formant transitions were removed between a word-initial fricative and the subsequent vowel, Pittman *et al.* (2002) found that performance of adult native English listeners with hearing loss deteriorated more than that of normal-hearing listeners on a fricative categorization task ([s, ʃ]) compared to when formant transitions were present. Listeners with age-related hearing loss may thus rely to some extent on formant transitions or other coarticulatory cues to distinguish between fricatives. Zeng and Turner (1990), on the other hand, concluded from their study on the recognition of four word-initial voiceless fricatives ([s, f, θ, ʃ]) that hearing-impaired native English listeners are able to use the fast and dynamic spectral information in formant transitions for fricative identification, though not as efficiently as normal-hearing listeners.

We present the results of two phonetic categorization experiments. In the first experiment, we investigated (1) whether older Dutch listeners use coarticulatory information in the vowel preceding the word-final fricative for the categorization of this word-final fricative to a larger extent than younger Dutch listeners; (2) whether high-frequency hearing loss differences within the group of older listeners (because older listeners generally show a larger variation in hearing sensitivity than younger adults) are associated with the use of this information. Additionally, we investigate the use of fricative noise information to assess the reliance on spectral cues for consonant identification by both age groups.

The stimuli consisted of Dutch minimal word pairs (e.g., *brief-bries*, “letter”-“breeze”) in which the final [f] or [s] was replaced by a manipulated, ambiguous sound from an [f]-[s] continuum. The manipulated sound was created such that the spectrum of the fricative noise was ambiguous and thus contained conflicting information about the fricative’s identity (see Sec. II B). Importantly, the preceding vowel and the original formant transitions leading into the fricative in the “source word” were left intact (this in contrast to the Pittman *et al.*, 2002 study where formant transitions were removed). Consequently, the coarticulatory information in the preceding vowel exposed the listener to varying degrees of conflicting cues (similar to Wagner *et al.*, 2006) regarding the fricative’s identity. The degree of conflict was largest when the two endpoints of the [f]-[s] continuum were combined with the conflicting source word. For example, if the *brie* taken from *brief* is concatenated to the most /s/-like step from the [f]-[s] continuum, the coarticulatory information conflicts maximally with the fricative noise. Conversely, the conflict is minimal when *brie* taken from *brief* is concatenated to the most /f/-like step from the [f]-[s] continuum. This manipulation allows us to investigate listeners’ reliance on fricative noise information and coarticulatory cues in the preceding vowel. Based on Wagner *et al.* (2006) and Pittman *et al.* (2002), we expect younger listeners to not use coarticulatory cues to differentiate [f] and [s], or at least to do so to a lesser extent than older listeners. Moreover, if age-related hearing loss causes older listeners to use coarticulatory cues to differentiate between [f] and [s], increasing reliance on coarticulatory cues is expected with increasing hearing loss among older adults. We will refer to this experiment as the “unfiltered condition” as all speech frequency information is available in the signal.

In the second phonetic categorization experiment (referred to as the “low-pass filtered condition”), we investigate whether the use of coarticulatory information becomes more pronounced when high-frequency spectral information is removed. Additionally, we investigate the use of fricative noise information to assess the deterioration in consonant categorization caused by the removal of high-frequency information. To that end, all stimuli were low-pass filtered and presented to a second group of younger and older listeners. Note that age-related high-frequency hearing loss is by no means simulated by the removal of all high-frequency information as in the filtered condition, as age-related hearing impairment encompasses more changes than just elevated thresholds, such as decreased temporal and frequency

resolution (Gordon-Salant *et al.*, 2010). Nevertheless, the removal of high-frequency spectral information allows us to investigate the reliance of younger and older listeners on high-frequency spectral information in the speech signal for the categorization of [f] versus [s]. If listeners use coarticulatory information from the preceding vowel to a larger extent in the low-pass filtered condition when compared to the unfiltered condition, this would suggest that listeners are able to immediately use coarticulatory information as a compensatory cue when fricative categorization becomes difficult.

## II. EXPERIMENTAL SET-UP

### A. Participants

Ninety-seven native Dutch participants were drawn from the MPI for Psycholinguistics subject pool and were paid for their participation. Two groups of “older” participants aged 60+ and two groups of “younger” participants were tested (see Table I), one group of each age category participated in either the filtered or the unfiltered condition. The group of older adults in the unfiltered condition was subsampled from Scharenborg and Janse (2012), such that the sample size, participants’ mean age, and mean hearing profile were matched as closely as possible with the newly recruited older listeners who participated in the low-pass filtered condition. Hearing sensitivity for all participants was assessed with a Maico ST20 portable screening audiometer (air conduction thresholds only) for octave frequencies from 250 Hz through 8 kHz. Mean high-frequency pure-tone averages (averaged over participants’ thresholds at 1, 2, and 4 kHz in their better hearing ear) for the four participant groups are listed in Table I. Figure 1 shows mean hearing sensitivity in terms of hearing thresholds from 250 to 8000 kHz for the older listeners in the unfiltered (solid lines) and filtered (dashed lines) conditions. The left ear is plotted in black; the right is plotted in grey. Neither the two younger [ $t(47) < 1, p > 0.1$ ] nor the two older [ $t(43) < 1, p > 0.1$ ] listener groups differed in their hearing sensitivity. Within each filtering condition, the younger and older listener groups differed in their hearing sensitivity [unfiltered condition group:  $t(31) < -12.10, p < 0.001$ ; filtered condition group:  $t(37) < -14.93, p < 0.001$ ]. Individual pure-tone average (over 1, 2, and 4 kHz in the better ear) of the older participants was entered as an index of hearing loss in the analyses. None of the participants wore hearing aids in daily life. No other hearing-related exclusion criteria were used.

TABLE I. The number of participants (number of males), age, and hearing information for each listener group.

	N (Male)	Age		Hearing (dB HL)	
		Mean (SD)	Mean (SD)	Mean (SD)	Range
Younger, full	24 (3)	21.3 (2.4)	0.0 (3.6)	-6.7-8.3	
Younger, filtered	25 (5)	21.8 (1.9)	-0.7 (3.9)	-6.7-10.0	
Older, unfiltered	24 (8)	72.0 (4.9)	23.8 (8.7)	10.0-36.7	
Older, filtered	24 (8)	72.2 (4.0)	23.1 (6.6)	13.3-36.7	

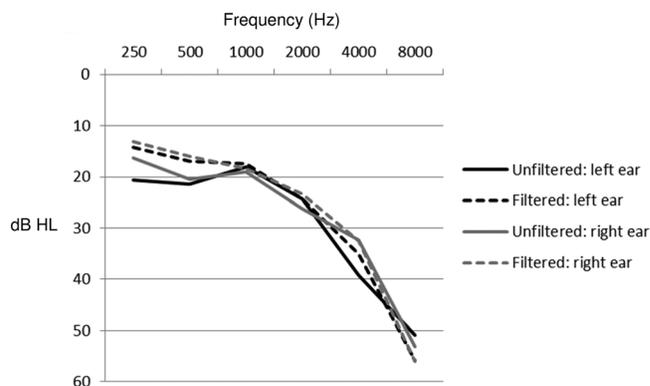


FIG. 1. Mean hearing sensitivity in terms of pure-tone hearing thresholds from 250 to 8000 kHz of the older listeners in the unfiltered (solid lines) and filtered conditions (dashed lines), separately for each ear.

### B. Materials and procedure

The test materials consisted of four minimal pairs of Dutch /f/- and /s/-final words: *brief-bries* (“letter”-“breeze”), *graf-gras* (“grave”-“grass”), *leef-lees* (“live”-“read”), *lof-los* (“praise”-“loose”). No filler words were included in the study. All words were produced in isolation by a female native speaker of Dutch and digitally recorded at 44.1 kHz in a sound-attenuated booth. Since it was not possible to find minimal pairs where both words had similar word frequency, word frequencies were balanced between the minimal pairs, i.e., two sets where the /s/-word had a higher frequency and two sets where the /f/-word had a higher frequency. Table II lists the minimal pairs and the individual word frequencies (obtained from SUBTLEX-NL, Keuleers *et al.*, 2010) of the words in the minimal pairs. Minimal pair was entered as a random predictor into the statistical models.

The procedure for creating the ambiguous [f/s] began with excision of the final fricative (i.e., [f] or [s]) of each word in each word pair (e.g., *brief* and *bries*) using PRAAT (Boersma and Weenink, 2005). The boundary between the vowel and the subsequent word-final fricative was set at a positive-going zero-crossing at the onset of the fricative noise on the basis of visual inspection of the wave form and spectrum. All excised [f]- and [s]-final sounds were zero-padded with 25 ms of silence at the onset and offset to allow for valid pitch estimation. Each word in the minimal pair then received the same stylized pitch contour (based on the naturally occurring pitch contour in the minimal pairs) using PRAAT. The excised and zero-padded [f] and [s] belonging to the same minimal pair (e.g., *brief*[f] and *brie*[s]) were then used in a time-aligned version of the morphing algorithm STRAIGHT (Kawahara *et al.*, 1999), which runs in MATLAB,

TABLE II. The word frequencies of the words in the minimal pairs. The word in the minimal pair with the highest frequency of the two is in bold.

	Frequency		Frequency
<i>brief</i>	<b>73.84</b>	<i>bries</i>	1.33
<i>graf</i>	<b>34.53</b>	<i>gras</i>	18.75
<i>leef</i>	46.67	<i>lees</i>	<b>61.15</b>
<i>lof</i>	3.84	<i>los</i>	<b>184.80</b>

to create an equally spaced 11-step continuum. STRAIGHT morphs sound files by first decomposing the input sound files into source parameters and spectral parameters (the source-filter model). Subsequently, pitch information is carefully removed, while keeping time-frequency information intact. Several parameters can then be carefully manipulated. In the creation of the [f]-[s] continuum in this study, the vocal tract and speaking rate parameters were carefully manipulated using interpolation, after which they were resynthesized. The resulting speech quality level is close to natural speech.

The resulting 11 morphed fricatives range from 100% [f] and 0% [s] (step 0) to 0% [f] and 100% [s] (step 10), where step 5 is a 50% [f] and 50% [s] blend. This procedure ensured that only the frication noise, which contains all the cues for differentiation between [f] and [s], was morphed, while leaving intact the formant transitions leading from the preceding vowel into the original fricative. Figure 2 shows the spectral envelopes of the fricative noises of the [f] taken from the original *lof* source word [Fig. 2(a)], the [f/s] sound [morph step 4, Fig. 2(b)], and the [s] taken from the original *los* source word [Fig. 2(c)]. The fricatives were taken from the word stimuli scaled to a mean intensity of 75 dB sound

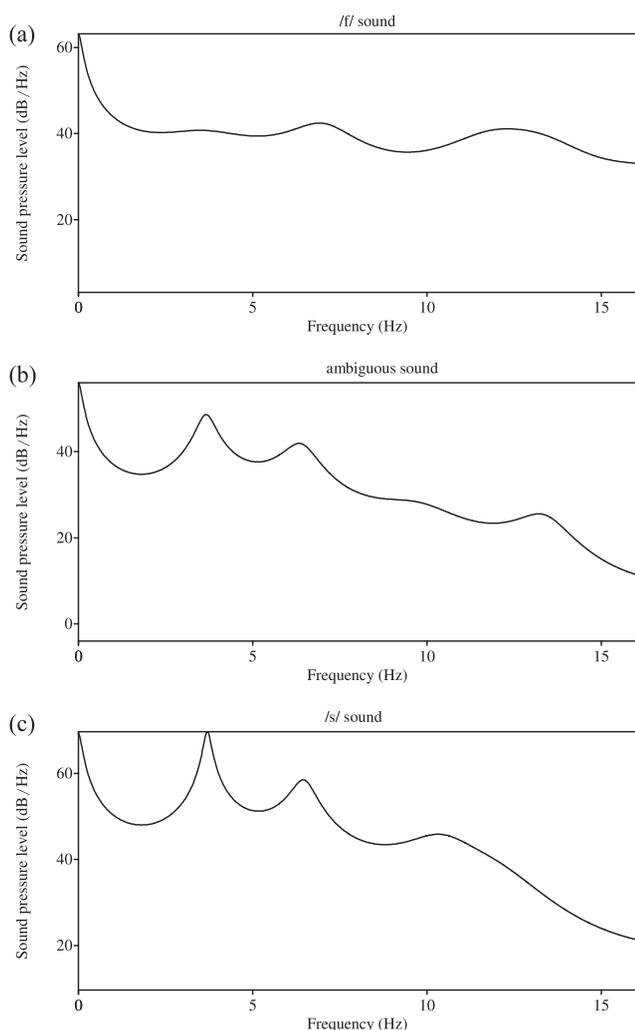


FIG. 2. Spectral envelopes of the fricative noises of (a) the original [f] sound from the *lof* source word, (b) the manipulated, ambiguous sound (morph step 4), and (c) the original [s] sound from the *los* source word.

pressure level (SPL). In the final step, the morphed fricatives were concatenated as final sounds to both items (i.e., to both *lo* from *lof* and to *lo* from *los*) in the minimal pair, resulting in 11 manipulated versions of each of the two words of a minimal pair.

The test continuum to be used in the phonetic categorization experiments was chosen on the basis of a pretest, which had as its goal to determine the most ambiguous [f/s] morph. The test continuum was subsequently chosen to sample the perception of different points along the continuum (excluding the endpoints) surrounding the most ambiguous morph in order to reduce response biases to either [f] or [s]. The pretest was carried out in the context of another experiment (Scharenborg *et al.*, 2015). The selection of the most ambiguous [f/s] sound was done separately for each of the four different vowels preceding final [f] and [s]. During the pretest, six [f]-[s] continuum steps (steps 1, 3, 4, 6, 7, and 9) for both words in each minimal pair were each presented six times binaurally via closed headphones. Ten older participants, who did not participate in the main experiments (though taken from the same pool as the main experiment participants), were tested individually in a sound-attenuated booth. The participants' task was to indicate whether they heard the /f/-final reading or the /s/-final reading of the word by pressing a button as quickly and as accurately as possible. To help the listeners, the two word choices were simultaneously presented for each trial, with the /f/-final word always printed on the left side of the screen and the /s/-final word on the right. The total proportions of /s/ responses to each of the tested morphs (averaged over all minimal pairs) were calculated. The most ambiguous morph step for each minimal pair was determined, which was step 4 for all minimal pairs. The test continuum was subsequently chosen with the most ambiguous morph (step 4) as its center; i.e., morph steps 1, 3, 4, 5, and 7. This resulted in 40 stimulus pairs (5 morph steps  $\times$  2 words per minimal pair  $\times$  4 minimal pairs). These "ambiguous" stimuli were then used in the phonetic categorization task.

Given our interest in whether listeners use coarticulatory cues such as formant transitions in the absence of high-frequency fricative noise, the formant transition distinguishing [s] from [f], i.e., F3 around 2200 Hz, had to remain intact. Accordingly, all stimuli were low-pass filtered at 2500 Hz using PRAAT. Note that, since the spectral information for [f] is distributed more uniformly over the spectrum, a small difference in spectral information between the low-pass filtered [f] and [s] noises remains present in the signal, mainly around 2000 Hz [compare Figs. 2(a) and 2(c)]. Given that the average hearing threshold at 2000 Hz was around 25 dB hearing level (HL) (see Fig. 1; ranging from 10 to 55 dB HL), the difference in energy between [f] and [s] in the filtered condition might have been inaudible to some of our participants.

During the actual experiments, participants were tested individually in a sound-treated booth. The word stimuli were presented binaurally over headphones at a fixed maximum level of 75 dB SPL. The two word choices were simultaneously presented for each trial, with the /f/-final word always presented on the left and the /s/-final word always on the

right of the screen. Keeping the location of the /f/-final and /s/-final words fixed allowed the participants to know the position of the two word choices even without looking at the screen. Participants were asked to press the button on the button box corresponding to the word they heard as fast and accurately as possible. They were not informed about the presence of manipulated sounds. The five manipulated, ambiguous items of each word in each minimal pair were each presented once per block (i.e., 40 items/block) and were newly randomized for each of a total of four blocks. The stimulus lists in the low-pass filtered condition were identical to those in the unfiltered condition, but the low-pass filtered version of the stimuli was used.

### III. RESULTS

Our first experiment (the unfiltered condition) sought to answer two questions: first, whether older Dutch listeners use coarticulatory information in the preceding vowel as a cue for word-final fricative categorization to a larger extent than younger Dutch listeners and, second, whether high-frequency hearing loss within the group of older listeners increases the use of this cue to fricative identity. Our second experiment (the filtered condition) sought to answer the question of whether the use of coarticulatory information becomes more pronounced when high-frequency spectral information is removed. Two analyses were carried out. In the first analysis, responses by the two age groups were compared in the two listening conditions. In the second analysis, the effect of hearing loss on acoustic cue use was investigated within the group of older listeners.

The analyses were carried out using generalized linear mixed-effect models (e.g., Baayen *et al.*, 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 *et al.*, 2013) to ensure that the effects found generalize across participants and minimal pairs. Inclusion of random slopes thus reduces the risk of type I errors.

The dependent variable in both analyses is whether the response is [f], coded as 0, or [s], coded as 1. We investigated the use of coarticulatory information in the preceding vowel as a cue to fricative categorization. Due to the manipulation of the stimuli, the fricative noise information is ambiguous while the original formant transitions are kept intact. If listeners are able to differentiate between /f/- and /s/-final source words, they thus do so by relying on the information in the preceding vowel. The use of coarticulatory cues for fricative categorization would then show itself as a significant effect of the fixed predictor Source Word (/f/-final vs /s/-final source word; the latter is on the intercept). In order to assess the reliance of listeners on fricative noise information for fricative categorization, Morph Step (i.e., morph steps 1, 3, 4, 5, 7) was entered as a fixed predictor in all analyses. The predictor Filtering indicates whether the stimuli are filtered ("1") or not ("0," on the intercept).

Moreover, we also investigate whether listeners become more sensitive to coarticulatory information, spectral noise information, or both over the course of the experiment. The fixed predictor Trial, as well as interactions between Trial and Morph Step and Trial and Source Word, was therefore entered in all analyses. To account for the possibility that the effect of coarticulatory information (i.e., Source Word) was modulated by information in the fricative noise, we also allowed for an interaction between Source Word and Morph Step. Additionally, in the first analysis, the predictor Age Group (i.e., younger vs older; the former are on the intercept) is added, while in the second analysis, Hearing Loss, as a numerical predictor, centralized to its mean, was added. We used dummy coding for binary fixed predictors (such as Source Word). Subject and Minimal pair were entered as random factors.

Each analysis started by building a model containing the random factors Subject and Minimal pair and all predictors and their interactions (as explained above). Subsequently, the data were analyzed by means of a backward stepwise selection procedure. First interactions and then predictors that were not significant at the 5% level were removed one-by-one from the model, with the least significant interaction or predictor always removed 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 than the previous model) by means of a likelihood ratio test with the `anova()` function in R. After the most parsimonious model containing only significant predictors and interactions in the fixed part was determined, the best maximal random slope structure was identified. Random by-subject slopes for Morph Step were added and tested through model comparisons in all analyses. In addition, random by-minimal-pair slopes for Hearing Loss were added in the analysis of the data of the older listeners. Changes in the random-slope structure were evaluated by means of the Akaike Information Criterion (AIC). The model with the lower AIC value and, therefore, better model fit was retained. Interactions and predictors in the fixed part that proved no longer to be significant (after changes in the random structure) were removed following the backward stepwise selection procedure. The final, best-fitting model thus only contains significant predictors and interactions. Only the statistically significant effects are reported. Furthermore, we report the absolute estimated values of the different  $\beta$ 's, which represent the relevance (effect size) of the different predictors for the estimation of the logit  $p$  (Chatterjee *et al.*, 2000).

Figure 3 shows the proportion of [s] responses for the five ambiguous [f/s] stimuli, averaged over the four test blocks for the four listener groups separately. Responses to stimuli originating from an /s/-final source word are indicated with a lower-case "s" and solid lines for the younger listeners and a capital "S" and dotted lines for the older listener group; likewise, /f/-final responses are indicated with "f" and "F" for stimuli originating from an /f/-final source. The results for the unfiltered and low-pass filtered conditions are visually depicted in Figs. 3(a) and 3(b), respectively.

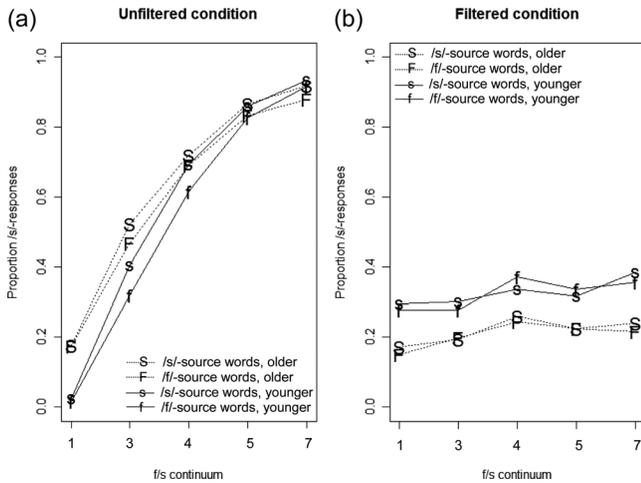


FIG. 3. The total proportion of [s] responses for the stimuli resulting from /s/-final source and /f/-final source words for younger (s vs f, solid lines) and older (S vs F, dotted lines) listeners in the unfiltered condition (a) as well as the low-pass filtered condition (b).

### A. Age group analysis

In the first analysis, the responses to the five ambiguous stimuli in the unfiltered and filtered condition for the younger and older listeners were compared. The start model contained the four possible four-way interactions of Source Word, Morph Step, Trial, Age Group, and Filtering. The parameter estimates in the best-fitting model of performance are shown in Table III under model A. If older Dutch listeners use coarticulatory cues for word-final fricative categorization to a larger extent than younger Dutch listeners, this would show itself as an interaction between Age Group and Source Word.

The best-fitting model (with the unfiltered condition on the intercept) showed an effect of Source Word: younger and older listeners gave significantly fewer [s] responses to /f/-final source words than to /s/-final source words [see also Fig. 3(a), where the s/S-lines are higher than the f/F-lines]. The lack of an interaction between Source Word and Age Group indicates that the difference between the proportions of [s] and [f] responses was similar for both age groups. This Source Word effect indicates that both younger and older Dutch listeners are able to use anticipatory coarticulatory cues in the speech signal for fricative categorization, and this effect is not modulated by information contained in the fricative noise (as there is no evidence of a Source Word by Morph Step interaction). Contrary to our prediction of an age-related difference in coarticulatory cue use, younger and older listeners use anticipatory coarticulatory cues to the same extent. Significantly more [s] responses were given over trials (see Trial effect in model A), and significantly more [s] responses were given to the more [s]-like morph steps on the continuum (see Morph Step effect in model A; in Fig. 3: the right hand side of each of the figures denotes more [s]-like stimuli; morph step 3 is on the intercept). Both the younger and older Dutch listeners thus also use fricative noise information for fricative categorization, although older listeners' categorization curves are shallower (i.e., they show smaller-sized Morph Step effects) compared to those of the

TABLE III. Fixed effect estimates for the best-fitting models of categorization performance. The dependent variable is the probability of [s] responses (logit). YA = younger adult data; OA = older adult data, n.s. = not significant.

Fixed effect	$\beta$	SE	$p <$
Model A: YA vs OA, unfiltered vs filtered (15 520 obs.):			
Intercept	0.5594	0.3488	n.s.
Source Word (/f/-word)	-0.3374	0.0674	0.001
Morph Step	1.1422	0.0811	0.001
Trial	0.0092	0.0030	0.01
Filtering	-1.5949	0.4710	0.001
Age Group	0.4966	0.6660	n.s.
Morph Step $\times$ Trial	0.0076	0.0021	0.001
Morph Step $\times$ Filtering	-1.0370	0.0924	0.001
Morph Step $\times$ Age Group	-0.1619	0.07862	0.05
Age Group $\times$ Filtering	-1.7277	0.6518	0.01
Source Word $\times$ Filtering	0.2902	0.0889	0.01
Trial $\times$ Filtering	-0.0094	0.0039	0.05
Morph Step $\times$ Trial $\times$ Filtering	-0.0071	0.0025	0.01
Model B: OA, unfiltered vs filtered (7680 obs.):			
Intercept	1.0331	0.4423	0.05
Source Word (/f/-word)	-0.1871	0.0663	0.01
Morph Step	1.1229	0.0727	0.001
Trial	0.0061	0.0029	0.05
Filtering	-3.4068	0.5697	0.001
Hearing Loss	-0.0603	0.0468	n.s.
Morph Step $\times$ Filtering	-1.0287	1.002	0.001
Morph Step $\times$ Hearing Loss	-0.0384	0.0082	0.001
Hearing Loss $\times$ Filtering	0.0673	0.0768	n.s.
Hearing Loss $\times$ Trial	-0.0010	0.0004	0.01
Morph Step $\times$ Filtering $\times$ Hearing Loss	0.0419	0.0133	0.01

younger listeners (Morph Step  $\times$  Age Group interaction in model A). The increase in [s] responses to more [s]-like stimuli increases over trials (see Morph Step  $\times$  Trial interaction in model A). In other words, over the course of the experiment, listeners showed an increase in use of spectral fricative noise information for fricative categorization, which may indicate an increased sensitivity to small differences in the fricative noise spectrum.

Removing the energy in the higher frequencies [compare Figs. 3(a) and 3(b)] clearly has a detrimental effect on fricative categorization. Both the younger and older listeners gave significantly fewer [s] responses to the low-pass filtered stimuli compared to the unfiltered condition stimuli (see Filtering effect in model A). Thus, as expected, filtering out the high-frequency spectral information removed important cues for [s] identification. The effect of filtering is significantly larger for the older adults compared to the younger adults (Filtering  $\times$  Age Group effect in model A), which can also be seen in Fig. 3: compare the dotted lines of the older listeners in Fig. 3(a), which lie mostly above the solid lines of the younger listeners, to the dotted lines in Fig. 3(b), which lie far below the solid lines.

As shown in Fig. 3, the effect of Source Word on fricative categorization is reduced after low-pass filtering the stimuli (Source Word  $\times$  Filtering in model A), but this interaction effect was particularly driven by the data of the younger adults. A separate analysis on only the data of the

younger adults in the filtered (rather than unfiltered) condition on the intercept showed that the simple effect of Source Word indeed disappeared. In the filtered condition, younger listeners thus do not use coarticulatory cues [see also Fig. 3(b) where the lines for the /f/- and /s/-source words are essentially superimposed]. As the analysis in the next section shows, the interaction between Source Word and Filtering was not found for the older adults (model B, cf. Sec. III B). The simple effect of Source Word for older adults in the filtered condition was confirmed to be still present through a separate analysis on the data of the older listeners only with the filtered (rather than the unfiltered) condition on the intercept [but note that this is difficult to tell from Fig. 3(b);  $\beta = -0.183$ ;  $SE = 0.066$ ;  $p < 0.01$ ]. This indicates that older Dutch listeners also use coarticulatory information in the vowel preceding the fricative to distinguish [s] from [f] when high-frequency spectral information is removed from the speech signal. Note, however, that the use of this coarticulatory information does not much help fricative categorization by the older listeners as they had a strong preference for [f] responses in the filtered condition. Older listeners thus seem to use coarticulatory information in the preceding vowel for fricative identification in both the unfiltered and filtered conditions.<sup>1</sup> Still, these results do not confirm our expectation that the removal of high-frequency spectral information would result in an increased reliance on coarticulatory cues for fricative categorization.

Returning to model A, the Morph Step effect is eliminated after the removal of high-frequency spectral information<sup>2</sup> (see interaction between Morph Step and Filtering in model A). Both younger and older listeners are thus no longer able to use fricative noise information for fricative categorization when the crucial high-frequency spectral information is removed. When the high-frequency spectral information is removed, younger and older listeners showed a reduced increase in [s] responses over trials compared to the unfiltered condition (Trial  $\times$  Filtering interaction in model A). Whereas stimuli were generally perceived as more [s]-like over trials in the unfiltered condition, this was less the case in the filtered condition. Moreover, the increase in sensitivity to small differences in the fricative noise spectrum observed in the unfiltered condition is less clear in the filtered condition (see the three-way interaction between Morph Step, Trial, and Filtering in model A).

The fixed, simultaneous presentation of the words of each minimal pair, with the /f/-final word always on the left and the /s/-final words always on the right bottom of the computer screen, may have possibly biased listeners' categorization behavior towards [f], as they would always read the /f/-final word first. However, model A showed that the intercept value (given in the first line) did not differ significantly from zero, suggesting that there was no bias for either response category for the condition mapped on the intercept (younger adults, most ambiguous Morph Step, unfiltered condition, source word = /s/, initial trials). The same analysis but with the /f/-final source word mapped on the intercept confirmed that listeners had no initial bias for either response category (the Trial effect may, however, indicate that listeners developed an [s]-bias over trials). Concerning possible

item biases, words within a minimal pair were either roughly equal in frequency, or one was clearly more frequent than the other (cf. Table II where the *lof-los* pair showed a frequency advantage for the /s/-word and the *brief-bries* pair showed a frequency advantage for the /f/-word), which is partially reflected in the random minimal pair effects. The *lof-los* pair ( $-0.080$ ) had a smaller negative random intercept than the *brief-bries* pair ( $-0.178$ ), indicating that the categorization data had to be adjusted depending on the minimal pair, and this pattern was also found for the subsequent analysis. The direction of the random intercepts for the other two minimal pairs, which were closer together in frequency, was not consistent across the two analyses. Last, the maximal random slope structure of model A included a participant random slope for Morph Step, indicating that the degree of increase in [s] responses differs over participants. Such individual differences in response to the morph steps could be due to auditory processing differences or dissimilar categorization biases.

## B. Hearing loss analysis

To investigate the role of hearing loss, we further analyzed the older adults' data.<sup>3</sup> If high-frequency hearing loss in older listeners increases the use of coarticulatory cues for fricative categorization, this would manifest itself as a significant interaction between Hearing Loss and Source Word. We also tested whether individual hearing loss interacted with Morph Step to investigate whether individuals with increased hearing loss would be less sensitive to acoustic differences between stimuli. The start model for the analysis of the older adults' data consisted of the four possible four-way interactions between Source Word, Morph Step, Trial, Filtering, and Hearing Loss. Table III, model B shows the parameter estimates in the best-fitting model of performance.

Similar results to the previous analysis were found regarding the simple effects of Source Word, Morph Step, Trial, and Filtering as well as for the interaction between Morph Step and Filtering. Importantly, there was no simple effect of Hearing Loss, nor was there a significant interaction between Hearing Loss and Source Word. There is therefore no indication that hearing loss modulates the use of the coarticulatory information contained in the source word. The use of coarticulatory information is thus unrelated to individual hearing loss in our sample. Hearing Loss did, however, interact with Morph Step (Morph Step  $\times$  Hearing Loss in model B): changes in the noise spectrum affected participants with poorer hearing less than better-hearing participants.

In other words, hearing loss makes participants less sensitive to acoustic differences between the morph steps. Hearing Loss also interacted with Trial: the general increase in [s] responses participants showed over trials was smaller the more hearing loss the participant had. Finally, the three-way interaction between Morph Step, Hearing Loss, and Filtering showed that the filtering effect on the use of the fricative noise information was less pronounced for those with more hearing loss. This is logical, as their use of the fricative noise information was poorer to begin with in the unfiltered

condition. The maximal random slope structure of model B included random effects for Subject and Minimal pair, a participant random slope for Morph Step, and a Minimal pair random slope for Hearing Loss indicating that the degree to which hearing loss predicted the proportion of [s] responses differed across minimal pairs. Note also that the intercept value in model B did differ significantly from zero, indicating that there was a slight general preference to respond with [s] in the older adults [see also the upward shift of the categorization lines for the older listeners compared to the younger listeners in Fig. 3(a)] in the condition mapped on the intercept (most ambiguous Morph Step, Source Word ending in /s/, and unfiltered condition). This [s] preference became marginally significant in the model B in which the Source Word ending in /f/ was mapped on the intercept (keeping everything else the same).

#### IV. GENERAL DISCUSSION

Age-related decline in hearing particularly affects the higher frequencies. Consequently, older listeners with hearing loss are less able to rely on spectral cues for the identification of fricatives that have their distinguishing noise information in those higher frequency regions. This research was inspired by the question whether listeners who suffer from age-related hearing loss have developed alternative strategies to correctly categorize or identify fricatives. The present study investigates the possibility that these listeners have learned to use other cues in the speech signal for fricative identification. In the first experiment, the use of coarticulatory information (e.g., formant transition information) in the vowel preceding the word-final fricative was investigated for word-final fricative categorization, comparing younger and older Dutch listeners. Moreover, within the group of older listeners, the effect of hearing loss on the use of coarticulatory information for fricative categorization was examined. The second experiment, also including younger and older listeners, investigated whether the use of coarticulatory cues became more pronounced when high-frequency spectral information was removed. In both experiments, listeners' use of fricative noise information for fricative categorization was investigated as well.

The experiments consisted of a phonetic categorization task in which listeners had to decide on the identity of a sound manipulated to fall somewhere between [f] and [s] in the context of Dutch minimal pairs (e.g., *brief* "letter" vs *bries* "breeze"). These manipulated sounds were created such that the spectrum of the fricative noise contained ambiguous information about the fricative's identity while the original formant transitions leading into the fricative in the "source word" were left intact. This manipulation allowed us to investigate listeners' reliance on fricative noise and coarticulatory information for word-final fricative categorization. The second experiment differed from the first experiment only in the low-pass filtering of all stimuli.

Our results showed that fricative noise information is the primary cue for Dutch fricative categorization. Manipulation of the fricative noise changed category identity from less than 20% [s] responses to less than 20% [f] responses (compare

Morph Step 1 and 5 in Fig. 3), with coarticulatory information in the preceding vowel acting as a less salient secondary cue. Our analyses showed that both the younger and the older Dutch listeners in our study used fricative noise information for fricative categorization and even increased its use over the course of the experiment. This suggests that the fricative noise information was a cue listeners became even more sensitive to over time. Hearing loss, however, impacted the use of the fricative noise information for fricative categorization, particularly for the recognition of [s]. Poorer-hearing listeners were less sensitive to acoustic differences between the morph steps than better-hearing listeners. Similarly, the filtering manipulation, which filtered out the high-frequency spectral information, eliminated the effect of fricative noise spectrum on categorization: both younger and older listeners were no longer able to use fricative noise information for fricative categorization when the crucial high-frequency spectral information was no longer available.

The crucial question was whether decreased use of the fricative noise spectrum can be traded off with increased reliance on other acoustic cues for fricative categorization and whether listeners are able to use other compensatory cues for the categorization of [f] and [s] when fricative categorization becomes difficult (i.e., upon the presentation of filtered speech). We therefore investigated whether listeners became more sensitive to coarticulatory information in the vowel preceding the fricative in experiment 2 (compared to experiment 1) and over the course of experiment 2. Alternatively, such changes in cue use may be more gradual and therefore take more time. If so, such cue tradeoffs might only be found in older adults with acquired hearing loss as they have built up more experience having limited access to fricative noise information.

Our results showed that when the full spectral information is presented to the listener, both younger and older Dutch listeners are able to differentiate between /f/- and /s/-final source words on the basis of coarticulatory information. These results are in contrast to Wagner *et al.* (2006), who showed that younger Dutch listeners are not affected by conflicting formant information for fricative categorization. These contrasting results could be due to the experimental paradigms used: a two-alternative forced choice fricative categorization task in the present study versus a phoneme-monitoring task in Wagner *et al.* (2006). As Wagner and colleagues have already pointed out, these differences in experimental paradigms may lead to differences in listening strategies with a stronger focus on perceptual/acoustic cues in the categorization task. Our findings that younger and older Dutch listeners can use coarticulatory cues extend the English-language results of Pittman *et al.* (2002) and Zeng and Turner (1990).

Contrary to our hypothesis, younger and older listeners use coarticulatory information to the same extent. Note that younger and older listeners not only differed in age but also in their hearing thresholds (cf. Table I for hearing differences between our younger and older listeners). However, among the older adults in our participant sample, hearing loss differences also did not lead to increased (nor decreased) use of coarticulatory cues for fricative categorization. Our results

contradict those reported by Pittman *et al.* (2002), who found that hearing-impaired listeners relied more on formant transition information than normal-hearing adults. This difference in findings may relate to a difference in experimental stimuli (word-initial in their study vs word-final fricatives in our study), experimental manipulation (Pittman *et al.* removed the formant transitions, where in our study they were kept intact), or participant sample. Regarding the latter, our older participant sample was still relatively good-hearing, at least compared to the hearing-impaired listeners in the Pittman *et al.* study. Moreover, the more subtle hearing loss differences in our study did not compare to the split between normal-hearing and poor-hearing participants in Pittman *et al.* Alternatively, in their phoneme monitoring experiment, Wagner *et al.* (2006) showed that young English listeners, unlike the Dutch listeners, were affected by misleading formant transitions for fricative identification. The differences in findings between the current study and that by Pittman and colleagues could thus be explained by the inherent difference in the importance of the coarticulatory cue for fricative categorization in Dutch and English.

Both the younger and older listeners became more sensitive to the acoustic differences in fricative noise over the course of the experiment (as shown by the Trial  $\times$  Morph Step interaction). They did not, however, become more sensitive to coarticulatory information over the course of the experiment. Moreover, importantly, the use of coarticulatory cues did not become more pronounced for either group when high-frequency information was filtered out and fricative categorization became difficult. After filtering, the use of coarticulatory information was even completely eliminated for the younger listeners, although the older adults still seemed to use it (cf. our discussion of what models A and B suggest in Sec. III). Regardless, these data do not provide any evidence that (older) listeners increase their use of coarticulatory information when high-frequency information is no longer available. Even though the use of coarticulatory information such as formant transitions for fricative categorization is a strategy employed in childhood, listeners do not seem to readily return to this secondary, and arguably weak, cue once they have learned to rely on spectral noise information for fricative categorization.

Our results show that both younger and older Dutch listeners are sensitive to fricative noise information as well as coarticulatory information in the preceding vowel in fricative categorization, with fricative noise being the stronger cue. Even though the coarticulatory cue distinguishes between [f] and [s] in a lower frequency range than fricative noise does, neither high-frequency hearing loss nor low-pass filtering made listeners rely more on the less salient coarticulatory cue. If information in the higher frequencies is less strongly available, Dutch listeners thus do not seem to learn to use coarticulatory cues in the speech signal for fricative identification; instead, fricative identification is simply impaired. Further research is required to determine whether listeners change their acoustic cue use if they receive more explicit instructions on which types of information to use, or if they have experienced more severe hearing loss over a longer period of time.

## ACKNOWLEDGMENTS

The research by O.S. was funded partly by the Max Planck International Research Network on Aging and partly by a Vidi-grant from the Netherlands Organisation for Scientific Research (NWO, Grant No. 276-89-003) and was carried out at the MPI for Psycholinguistics, The Netherlands. E.J. was supported by a separate Vidi-grant from NWO (Grant No. 276-75-009). A.W. was funded by the Max Planck Society. We thank the student assistants of the MPI for Psycholinguistics for assistance in running the experiments and M. Witteman for recording the stimuli.

<sup>1</sup>Note that there was no three-way interaction between Source Word, Filtering, and Age Group in model A.

<sup>2</sup>The simple effect of Morph Step disappeared for both the younger and older adults when the filtered condition was mapped on the intercept.

<sup>3</sup>An identical analysis on the younger listeners' data revealed no effect of Hearing Loss, confirming our earlier observation of the lesser variation in hearing sensitivity in our younger listener group compared to our older listener group.

- Apoux, F., and Bacon, S. P. (2004). "Relative importance of temporal information in various frequency regions for consonant identification in quiet and in noise," *J. Acoust. Soc. Am.* **116**, 1671–1680.
- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). "Mixed-effects modeling with crossed random effects for subjects and items," *J. Mem. Lang.* **59**, 390–412.
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *J. Mem. Lang.* **68**, 255–278.
- Boersma, P., and Weenink, D. (2005). "Praat. Doing phonetics by computer (Version 5.1)." <http://www.sprache.de/~boersma/>
- Chatterjee, S., Hadi, A. S., and Price, B. (2000). *Regression Analysis by Example* (John Wiley & Sons, New York).
- Gelfand, S. A., Piper, N., and Silman, S. (1986). "Consonant recognition in quiet and in noise with aging among normal hearing listeners," *J. Acoust. Soc. Am.* **80**(6), 1589–1598.
- Gerrits, P. A. M. (2001). "The categorisation of speech sounds by adults and children," Ph.D. thesis, University of Utrecht, the Netherlands.
- Gordon-Salant, S., Frisina, R. D., Popper, A., and Fay, D. (2010). *The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis* (Springer, Berlin, Germany).
- Hedrick, M. S., and Ohde, R. N. (1993). "Effect of the relative amplitude of friction on the perception of place of articulation," *J. Acoust. Soc. Am.* **94**, 2005–2026.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**, 1252–1263.
- Kawahara, H., Masuda-Katsuse, I., and 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 Commun.* **27**, 187–207.
- Keuleers, E., Brysbaert, M., and New, B. (2010). "SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles," *Behav. Res. Methods* **42**(3), 643–650.
- Lippmann, R. P. (1996). "Accurate consonant perception without mid-frequency speech energy," *IEEE Trans. Speech Audio Process.* **4**, 66–69.
- Maniwa, K., Jongman, A., and Wade, T. W. (2008). "Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners," *J. Acoust. Soc. Am.* **123**, 1114–1125.
- Miller, G. A., and Nicely, P. (1955). "An analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Am.* **27**, 338–352.
- Monson, B. B., Hunter, E. J., Lotto, A. J., and Story, B. H. (2014). "The perceptual significance of high-frequency energy in the human voice," *Front. Psychol.: Auditory Cognitive Neurosci.* **5**, 587.
- Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2010). "Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task," *J. Acoust. Soc. Am.* **128**, 360–371.

- Moore, B. C. J., and Tan, C. T. (2003). "Perceived naturalness of spectrally distorted speech and music," *J. Acoust. Soc. Am.* **114**, 408–419.
- Nittrouer, S. (1992). "Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries," *J. Phonetics* **20**, 351–382.
- Nittrouer, S. (2002). "Learning to perceive speech: How fricative perception changes, and how it stays the same," *J. Acoust. Soc. Am.* **112**, 711–719.
- Pittman, A. L., Stelmachowicz, P. G., Lewis, D. E., and Hoover, B. M. (2002). "Influence of hearing loss on the perceptual strategies of children and adults," *J. Speech Lang. Hear Res.* **45**(6), 1276–84.
- R development core team (2011). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna.
- Rietveld, A. C. M., and van Heuven, V. J. (1997). *Algemene Fonetiek* (Dick Coutinho, Bussum, the Netherlands).
- Scharenborg, O., and Janse, E. (2012). "Hearing loss and the use of acoustic cues in phonetic categorisation of fricatives," in *Proceedings of Interspeech*, Portland, OR.
- Scharenborg, O., Weber, A., and Janse, E. (2015). "The role of attentional abilities in lexically-guided perceptual learning by older listeners," *Atten. Percept. Psychophys.* **77**(2), 493–507.
- Sher, A. E., and Owens, E. (1974). "Consonant confusions associated with hearing loss above 2000 Hz," *J. Speech Hear. Res.* **17**, 669–681.
- Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., and Lewis, D. E. (2001). "Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults," *J. Acoust. Soc. Am.* **110**, 2183–2190.
- Wagner, A. (2013). "Cross-language similarities and differences in the uptake of place information," *J. Acoust. Soc. Am.* **133**, 4256–4267.
- Wagner, A., Ernestus, M., and Cutler, A. (2006). "Formant transitions in fricative identification: The role of native fricative inventory," *J. Acoust. Soc. Am.* **120**, 2267–2277.
- Whalen, D. H. (1981). "Effects of vocalic formant transitions and vowel quality on the English [s]-[ʃ] boundary," *J. Acoust. Soc. Am.* **69**, 275–282.
- Zeng, F. G., and Turner, C. W. (1990). "Recognition of voiceless fricatives by normal and hearing-impaired subjects," *J. Speech and Hear Res.* **33**, 440–449.