# Hearing Loss and the Use of Acoustic Cues in Phonetic Categorisation of Fricatives

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

[1] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[2] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands
[3] Centre for Language Studies, Radboud University Nijmegen, The Netherlands
{Odette.Scharenborg, Esther.Janse}@mpi.nl

## Abstract

Aging often affects sensitivity to the higher frequencies, which results in the loss of sensitivity to phonetic detail in speech. Hearing loss may therefore interfere with the categorisation of two consonants that have most information to differentiate between them in those higher frequencies and less in the lower frequencies, e.g., /f/ and /s/. We investigate two acoustic cues, i.e., formant transitions and fricative intensity, that older listeners might use to differentiate between /f/ and /s/. The results of two phonetic categorisation tasks on 38 older listeners (aged 60+) with varying degrees of hearing loss indicate that older listeners seem to use formant transitions as a cue to distinguish /s/ from /f/. Moreover, this ability is not impacted by hearing loss. On the other hand, listeners with increased hearing loss seem to rely more on intensity for fricative identification. Thus, progressive hearing loss may lead to gradual changes in perceptual cue weighting.

**Index Terms**: fricative perception, aging, hearing loss, acoustic cues.

## 1. Introduction

Age-related declines in hearing particularly affect sensitivity to the higher frequencies (the higher the frequency, the greater the age-related sensitivity loss), which results in the loss of sensitivity to phonetic detail. One would therefore expect that age-related hearing loss (calculated as the average over participants' hearing thresholds at 1, 2 and 4 kHz in their better ear) interferes with the discrimination between two consonants that have most information to differentiate between the two consonants in those higher frequencies and less in the lower frequencies. An example of such a consonant contrast is /f/ vs. /s/. [s] has more energy in the higher frequencies, concentrating around 5500 Hz, while [f] has a flatter spectrum, with the energy distributed more uniformly over the spectrum [1]. This information in the higher frequencies may be less strongly available to many older listeners. Due to this loss of sensitivity to the higher frequencies, listeners with hearing loss may no longer be able to rely on the 'normal' perceptual strategies for distinguishing /f/ from /s/. One possibility is that listeners with hearing loss start to use other cues in the speech signal to differentiate between /f/ and /s/.

One such acoustic cue could be formant transitions. Pittman and colleagues [2] found that listeners with hearing loss performed worse on a fricative categorisation task (/s/ or /ʃ/) when formant transitions were removed between the word-initial fricative and the subsequent vowel (/ɑ/ or /æ/) compared to when these formant transitions were present. This seems to suggest to listeners with hearing loss (acquired later in life) use formant transitions to distinguish between fricatives. On the other hand, Zeng and Turner concluded from their study on the recognition of four word-initial voiceless fricatives (/s, f, θ, ʃ/) that hearing-impaired listeners are able to use the fast and dynamic spectral information in formant transitions for fricative identification in some cases but not as efficiently as normal-hearing listeners [3]. Note however, that formant transitions (in Dutch) from the preceding vowel into the following fricative are fairly similar for /f/ and /s/, with only a small difference in slope of F3 around 2200 Hz (a rise into a following /s/ and flat into a following /f/) [1]. Importantly, this distinctive information is represented at lower frequencies than that in the fricative noise itself.

Research suggests that there may be cross-linguistic differences in the use of formant transitions for fricative identification as a function of spectral similarity in the native language's fricative inventory [4]. Research on Dutch listeners (Dutch lacking a dental fricative) has shown that "vowel transitions do not contain perceptually relevant information about adjacent fricatives in Dutch" [5], p.79). Likewise, Wagner et al. [4] showed that Dutch listeners were not affected by misleading formant transitions for fricative identification of Spanish stimuli (Spanish having labiodental, dental and alveolar place of articulation fricatives), while Spanish listeners listening to the same Spanish stimuli were. Normal-hearing (young) Dutch adults thus normally do not seem to use formant transitions for fricative identification. However, it is possible that in the face of deteriorating hearing, listeners with hearing loss start to use formant transitions for fricative identification.

A second possible cue for fricative identification is intensity. /f/ normally has a lower intensity than /s/ (see also below). It might be the case that listeners with hearing loss use intensity as a cue for /f/-/s/ identification, so that 'soft' fricatives with a lower intensity are identified as /f/ and 'louder' fricative intensities as /s/.

The question addressed in this study is: which acoustic cues do older listeners use to differentiate between /f/ and /s/ as a function of their hearing loss? We investigate two possible cues: formant transitions (Experiment 1) and intensity (Experiment 2). Both experiments consist of a phonetic categorisation task.

In order to investigate what cues older listeners use to differentiate between /f/ and /s/, four Dutch minimal pairs of /f/- and /s/-final words were used. In Experiment 1, listeners were confronted with a range of ambiguous sounds from the [f]-[s]-continuum appearing as the final sound of both words of the four minimal pairs, and were asked to decide whether the final sound was /f/ or /s/. The ambiguous sounds were created such that the

spectrum of the final fricative contained conflicting information about the identity of the fricative, as the original formant transitions were left intact. It is to be expected that if hearing loss makes listeners use formant transitions to differentiate between /f/ and /s/, there would be more /s/-responses to /s/-source words and more /f/-responses to /f/-source words for those with poorer hearing. On the other hand, since the difference in formant transitions is mainly to be found in F3 around 2200 Hz, where age-related hearing loss also already has its effect, it is also possible that hearing loss does not increase the reliance on formant transitions to differentiate between /f/ and /s/.

In Experiment 2, the minimal pairs ended in natural /f/ and /s/ but now the intensity of the final fricative noise was changed from 44 dB to 56 dB (relative to 70 dB for the vowel portion). Participants again were asked whether the final sound was an /f/ or an /s/. The range of intensities was chosen on the basis of the intensity of natural occurring final /f/'s and /s/'s in the stimulus set. Note that in this case, the spectra of the final fricatives as well as matching formant transition information are available to the listener. If hearing impairment makes listeners rely more strongly on intensity to determine the identity of the final fricative, we would expect more /f/-responses for the lower intensities and more /s/-responses for the higher intensities particularly for those with more hearing loss.

# 2. Experimental set-up

## 2.1. Participants

Thirty-eight participants aged 60+ (15 M; mean age: 72.4; SD: 6.9) and native Dutch speakers were drawn from the MPI for Psycholinguistics subject pool and were paid for their participation. None of them wore hearing aids. Hearing sensitivity was assessed (in the context of other experiments, not reported here) with a Maico ST20 portable audiometer (air conduction thresholds only) for octave frequencies from 250 Hz through 8 kHz. Mean pure-tone average (averaged over participants' thresholds at 1, 2, and 4 kHz in their better ear) was 26.0 dB HL (SD=11.7).

## 2.2. Materials

The four minimal pairs of /f/- and /s/-final words were: *brief* - *bries* ('letter' - 'breeze'), *graf* - *gras* ('grave' - 'grass'), *leef* - *lees* ('live' - 'read'), *lof* - *los* ('praise' - 'loose'). All words were produced in isolation by a female native speaker of Dutch and digitally recorded in a sound-attenuated booth at 44 kHz.

### *2.2.1. Experiment 1: Formant transitions*

The ambiguous sounds used in Experiment 1 were created as follows. For each of the four minimal pairs, the final fricative was excised and zero-padded with 25 ms of silence at onset and offset to allow valid pitch estimation. Subsequently, each word received the same stylised pitch contour (based on the naturally occurring pitch contour of the words in the minimal pairs) using Praat. Next, the excised /f/ and /s/ belonging to the same minimal pair were morphed to create an equally-spaced 11-step continuum using STRAIGHT [6] in Matlab. The ambiguous fricatives were then concatenated as final sounds to both the /f/- and /s/-final source words. This procedure was followed to ensure that formant transitions were kept as natural as possible. Five versions of the stimuli with ambiguous [f/s] were subsequently used in Experiment 1 (i.e., steps 1, 3, 4, 5, 7; note that the ambiguous sounds were pretested in the context of another experiment [7]).

Figure 1 shows an example of the result of the morphing and concatenation procedure. The top panel shows the spectrogram and formant structure of *brief* with a natural final [f]; the second and third panels show the spectrogram and formant structure of *brie[f/s]* with the ambiguous final [f/s] (in this case step 5 from the continuum) concatenated to the source words *brief* and *bries*, respectively; and the bottom panel shows the spectrogram and formant structure of *bries* with a natural final [s].

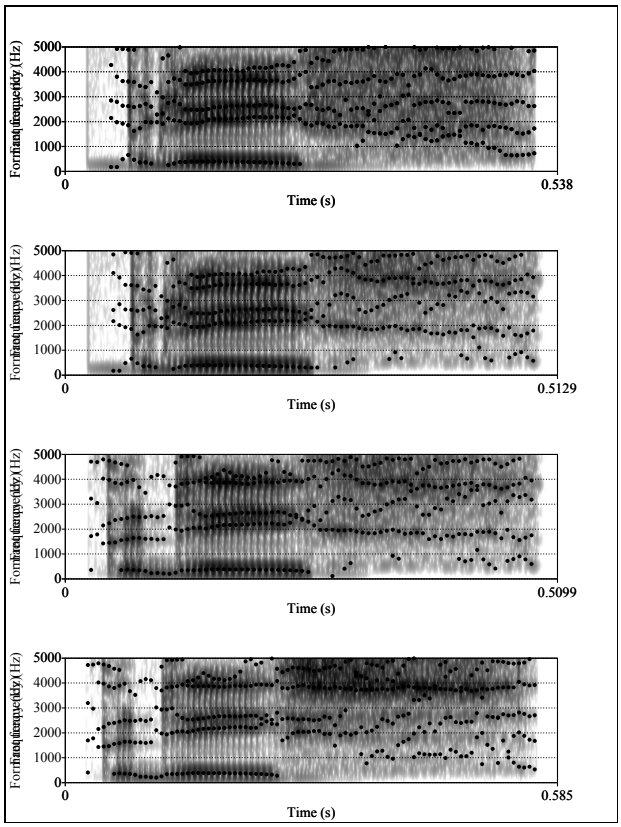


*Figure 1. The spectrogram and formant tracks of: top panel – 'brief' with a natural final [f]; second panel – brie[f/s] from the source word 'brief' with the ambiguous final [f/s] (in this case step 5 from the continuum); third panel – brie[f/s] from the source word 'bries' with the ambiguous final [f/s] (step 5 from the continuum); bottom panel – 'bries' with a natural final [s].*

### *2.2.2. Experiment 2: Intensity*

To investigate the role of intensity on the identification of final fricatives, versions of the test items of the four minimal pairs used in Experiment 1 were created in which the intensity of the noise of the final fricative was varied. For each test item of these four minimal pairs, the final natural fricative was excised, and the intensity of the word onset (e.g., *brie*) was set at 70 dB. Seven versions of each of the final fricatives were subsequently created by varying their intensity between 44 dB and 56 dB in increments of 2 dB. This range was based on the intensities of the natural /f/ and /s/'s in these eight stimuli. Subsequently, the

intensity-modulated final fricatives were concatenated as final sounds to their corresponding /f/- or /s/-final source word. The resulting stimuli were natural words, but with a final fricative that was intensity-modulated.

### 2.3. Procedure

In both experiments, the participants were tested individually in a sound-treated booth. The stimuli were presented binaurally over closed headphones at a fixed maximum level. To aid the listeners, for each auditory stimulus, both words of its minimal pair were presented on the screen. The /f/-final word was always presented on the bottom-left and the /s/-final word always on the bottom-right of the screen. Participants were asked to press the button corresponding to the word they heard as fast and accurately as possible. They were not informed about the presence of ambiguous or intensity-modulated sounds.

In Experiment 1, the five ambiguous items of each word in each minimal pair were each presented once per block (i.e., 40 items/block), and were newly randomised for each of a total of four blocks (160 items in total). In Experiment 2, the seven intensity-modulated items of each test item in each minimal pair were each presented once per block (i.e., 56 items/block), and were newly randomised for each of a total of two blocks.

## 3. Results and discussion

### 3.1. Experiment 1: Formant transitions

Due to failure of the experimental software, the results of one participant were not recorded. The phonetic categorisation data were analysed using generalised linear mixed-effects models. The results presented here were obtained with the best-fitting model (after model comparisons). Figure 2 shows the proportion of /s/-responses for the five ambiguous [f/s] stimuli, averaged over the four test blocks. In order to investigate whether listeners use formant transitions in the face of ambiguous final fricatives, the data are split into two groups, i.e., responses to the stimuli that originated from an /s/-final source word (indicated with 'S') and responses to the stimuli that originated from an /f/-final source word (indicated with 'F'). The research question addressed here is whether high-frequency hearing loss relates to use of formant transition information for fricative categorisation. This would indeed be the case if we find an interaction between hearing loss and source word.

As can be seen in Figure 2, there is an effect of source word on phonetic categorisation. Significantly fewer /s/-responses were given to /f/-final source words than to /s/-final source words ($\beta$ = -.2056, $SE$ = .0735, $p$ < .01). This result suggests that (at least) Dutch older listeners are able to use cues in the speech signal other than those found in final fricative to determine the identity of the final fricative when fricative information is ambiguous, and these cues are likely to be formant transitions. Moreover, as expected, there is an increase in /s/-responses for more /s/-like stimuli (the higher steps on the continuum; $\beta$ = .9879, $SE$ = .0307, $p$ < .001).

There is a general effect of hearing loss, with fewer /s/-responses with increasing hearing loss ($\beta$ = -.0591, $SE$ = .0260, $p$ < .05), the latter being particularly the case for more /s/-like stimuli ($\beta$ = -.0465, $SE$ = .0029, $p$ < .001). These results indicate that people with increasing hearing loss have more trouble recognising /s/. This is in line with our hypothesis that when a listener has trouble hearing the higher frequencies, this will particularly impact the recognition of sounds that have their distinguishing cues in the higher frequencies (i.e., compare the spectrum of [s] with more energy around 5500 Hz, with that of the [f] which has the energy distributed more uniformly over the spectrum). Hearing loss thus impacts the use of spectral information for distinguishing /s/ from /f/. However, there is no indication that hearing loss interacts with the use of the information contained in the source word; showing that listeners with varying degrees of hearing loss are equally well able to use formant transitions to distinguish /s/ from /f/.

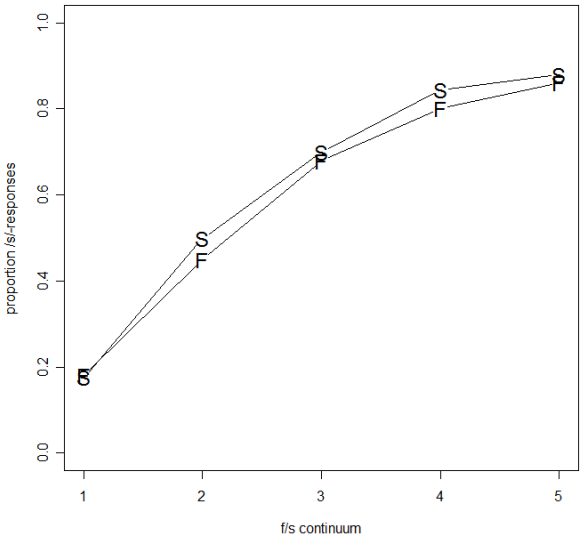


*Figure 2. The total proportion of /s/ responses for the stimuli resulting from /s/-final source words (S) and for stimuli resulting from the /f/-final source words (F).*

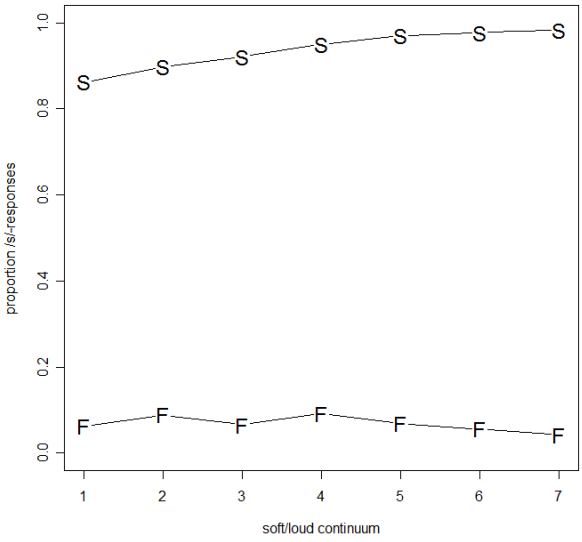


*Figure 3. The total proportion of /s/ responses to the intensity-modulated stimuli for the /s/-final words (S) and for the /f/-final words (F).*

### 3.2. Experiment 2: Intensity

The phonetic categorisation data of Experiment 2 were analysed using generalised linear mixed-effects models. The results presented here were obtained with the best-fitting model (after model comparisons). Figure 3 shows the proportion of /s/-responses for the seven intensity-modulated stimuli, averaged over the two test blocks. In order to investigate whether listeners use intensity to distinguish /f/ from /s/, the data are split into two groups, i.e., the responses to the /s/-final stimuli (indicated with 'S') and the responses to the /f/-final stimuli (indicated with 'F'). The research question addressed here is whether high-frequency hearing loss relates to the use of intensity information for fricative categorisation.

As Figure 3 clearly shows, not surprisingly, there are significantly fewer /s/-responses to /f/-final words than to /s/-final words ($\beta$ = -7.6817, *SE* = .2790, $p < .001$). There is an effect of hearing loss: with increased hearing loss, there are significantly fewer /s/-responses ($\beta$ = -.1411, *SE* = .0202, $p < .001$); although this is less so for /f/-final words ($\beta$ = .2990, *SE* = .0224, $p < .001$). Like was found in Experiment 1, these results suggest that listeners with high-frequency hearing loss have more problems identifying /s/ than identifying /f/, as one would expect on the basis of the spectra for /s/ and /f/. Interestingly, with increasing hearing loss, there are more /s/-responses with increasing intensity ($\beta$ = .0138, *SE* = .0041, $p < .001$); although this is (unsurprisingly) less so for /f/-final words ($\beta$ = .0167, *SE* = .0056, $p < .005$). This result indicates that hearing loss makes listeners rely more strongly on intensity as a cue to decide whether the heard phoneme is an /s/ or /f/, with a larger intensity leading to relatively more /s/-responses.

## 4. General discussion and conclusions

In this paper, we investigate the question which cues older listeners use when trying to differentiate between two consonants, /f/ and /s/, which have their differentiating information mostly in the higher frequency regions. In particular, we investigate how age-related high-frequency hearing loss may influence which cues are used. In two experiments, we investigate the use of formant transitions (Experiment 1) and fricative noise intensity (Experiment 2). Both experiments consist of a self-paced phonetic categorisation task in which participants have to indicate whether they have heard the /f/- or /s/- interpretation of four different minimal Dutch word pairs. In Experiment 1, the critical final fricative noise was replaced by an ambiguous noise sound in between /f/ and /s/. In Experiment 2, the intensity of the critical (natural) final fricative was either increased or decreased compared to its normal intensity.

In line with findings by [2] and [3], Experiment 1 shows that Dutch older listeners seem to use formant transitions, at least when the fricative spectra are ambiguous and do not contain unambiguous information about the identity of the fricative. However, unlike [2] and [3], our results seem to suggest that hearing loss does not interfere with the ability to use formant transitions for fricative identification. The difference between our results and [2] and [3] might be due to different reasons. First, we used word-final /f/ and /s/, whereas [2] and [3] used word-initial /s/ and /ʃ/, which are spectrally more similar. Second, it might be that the hearing loss suffered by the participants in these three studies was different. This is however difficult to assess. Third, we tested native Dutch listeners, while [2] and [3] tested native English listeners. It might be that different cues have a different weight or role in fricative identification in different languages, regardless of hearing loss, see e.g., [4].

In Experiment 2, the use of intensity as a cue to distinguish /f/ from /s/ was investigated. The results show that with increased hearing loss, listeners' perception of /s/ deteriorates; however, this was less so for /s/ stimuli with higher intensities. These results show that listeners when faced with increased hearing loss use intensity of the fricative to determine the identity of that fricative. We should note, however, that the effects of intensity are fairly small (see also Figure 3), and are mainly driven by the listeners with hearing loss.

In our study, we did not test young normal-hearing Dutch adults (yet) on the same stimulus set. We therefore do not know whether they use formant transitions and intensity to differentiate between /f/ and /s/, as was found for the older listeners. Nevertheless, Wagner et al. showed that normal-hearing young Dutch adults normally do not seem to use formant transitions for fricative identification [4]. A question that therefore arises is the cause of this (apparent) change in perceptual cue weighting strategy: do older listeners use the formant transitions and intensity cues due to progressing age-related hearing loss or is it a skill every listener of Dutch can use immediately when the need arises? This question will be investigated in a follow-up experiment with young Dutch normal-hearing adults.

To conclude, older listeners seem to use formant transitions as a cue to distinguish word-final /s/ from word-final /f/. This ability is not impacted by hearing loss. Moreover, listeners with increased hearing loss seem to rely more on intensity for fricative identification. Thus, progressive hearing loss may lead to gradual changes in perceptual cue weighting.

## 5. Acknowledgements

The research by Odette Scharenborg is sponsored by the Max Planck International Research Network on Aging. The research by Esther Janse is supported by a Vidi-grant from the Netherlands Organisation for Scientific Research (NWO). We thank our student-assistants for assistance in preparing and running these experiments, and Marijt Witteman for recording the stimuli.

## 6. References

[1] Rietveld, A.C.M., van Heuven, V.J. "Algemene Fonetiek", Bussum: Dick Coutinho, 1997.

[2] Pittman, A.L., Stelmachowicz, P.G., Lewis, D.E., Hoover, B.M. "Influence of hearing loss on the perceptual strategies of children and adults", J Speech Lang Hear Res. 45(6):1276-84, 2002.

[3] Zeng, F.G., Turner, C.W. "Recognition of voiceless fricatives by normal and hearing-impaird subjects", J Speech and Hear Res., 33:440-449, 1990.

[4] Wagner, A., Ernestus, M., Cutler, A. "Formant transitions in fricative identification: The role of native fricative inventory", JASA, 120:2267-2277, 2006.

[5] Klaassen-Don, L.E.O. "The influence of vowels on the perception of consonants", Doctoral dissertation, Leiden University, The Netherlands, (unpublished), 1983.

[6] Kawahara, H., Masuda-Katsuse, I., Cheveigne, A. "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, 1999.

[7] Scharenborg, O., Janse, E., Weber, A. "Perceptual learning of /f/-/s/ by older listeners", Proceedings of Interspeech, Portland, OR, 2012.