Changes in the Role of Intensity as a Cue for Fricative Categorisation

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

[1] Centre for Language Studies and Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

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

{O.Scharenborg, E.Janse}@let.ru.nl

Abstract

Older listeners with high-frequency hearing loss rely more on intensity for categorisation of /s/ than normal-hearing older listeners. This study addresses the question whether this increased reliance comes about immediately when the need arises, i.e., in the face of a spectrally-degraded signal. A phonetic categorisation task was carried out using intensity-modulated fricatives in a clean and a low-pass filtered condition with two younger and two older listener groups. When high-frequency information was removed from the speech signal, younger listeners started using intensity as a cue. The older adults on the other hand, when presented with the low-pass filtered speech, did not rely on intensity differences for fricative identification. These results suggest that the reliance on intensity shown by the older hearing-impaired adults may have been acquired only gradually with longer exposure to a degraded speech signal.

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

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. Information about place of articulation of consonants is particularly vulnerable to hearing loss [1],[2],[3]. An example of a consonant contrast for which high-frequency information is crucial is /s/-/f/ [4]. While [s] has its energy concentrated around 5500 Hz, for [f] the energy is distributed more uniformly over the spectrum [4]. Moreover, [f] normally has a lower intensity than [s] [5]. Listeners with age-related hearing loss may therefore no longer be able to rely on the 'normal' perceptual strategies for distinguishing /f/ from /s/. One possibility is that they no longer can distinguish between /f/ and /s/. Another possibility is that listeners with hearing loss start to use other cues in the speech signal for fricative categorisation [1].

Indeed, [6] found that older native English listeners with hearing loss performed worse on a fricative categorisation task (/s/, /ʃ/) when formant transitions were removed between the word-initial fricative and the subsequent vowel (/ɑ/, /æ/) compared to when these formant transitions were present. This seems to suggest to listeners with acquired age-related hearing loss use formant transitions to distinguish between fricatives. On the other hand, [7] concluded from their study on the recognition of four word-initial voiceless fricatives (/s, f, θ, ʃ/) that native English 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. Hedrick and Younger [1] found that, in discriminating /s/ and /ʃ/, normal-hearing listeners gave nearly equal weight to transition and relative amplitude information, whereas listeners with sensorineural hearing loss gave more weight to relative amplitude.

In a previous study [5], we investigated how age-related high-frequency hearing loss may influence which acoustic cues, more specifically, formant transitions and intensity, are used for the categorisation of /f/ vs. /s/ in Dutch [5]. The results showed that older listeners with hearing loss seem to rely more on intensity for fricative identification than older listeners with normal hearing.

An open question, which this study addresses, is whether this difference in cue use is due to progressive age-related hearing loss or whether the increased reliance on intensity for fricative identification may come about immediately when the need arises. We try to answer this question by investigating younger and older listeners' use of intensity as a cue to fricative categorisation in clean speech ('clean condition') and when high-frequency spectral information is removed ('low-pass filtered condition'). The results of the younger listeners will be compared to those of the older listeners to investigate whether younger listeners use intensity as a cue, and whether its use differs from older listeners.

It is important to note that age-related high-frequency hearing loss is by no means simulated by the removal of all high-frequency information as is done in the filtered condition, as age-related hearing impairment involves more than elevated thresholds, such as decreased temporal and frequency resolution [8]. 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 categorisation of /f/ vs. /s/, and thus helps answer our research question.

We used the set-up from [5], which consisted of a phonetic categorisation task with stimuli consisting of four Dutch minimal pairs of /f/-final and /s/-final words (such as *brief/bries* ('letter', 'breeze'). The intensity of the final fricative noise was adapted to range from 44 dB to 56 dB (relative to a mean of 70 dB for the vowel portion). Participants were asked whether the final sound was /f/ or /s/. Note that the spectra of the final fricatives as well as intact formant transition information in the preceding vowel are available to the listener. To investigate the contribution of spectral information contained in the higher frequencies for the recognition of /s/ vs. /f/, we low-pass filtered all stimuli.

The use of intensity for fricative categorisation manifests itself as an increase in /f/ responses for the lower intensities and an increase in /s/ responses for the higher intensities, irrespective of the identity of the final fricative. If such a pattern is found for the listeners in the low-pass filtered condition but not in the clean condition, this would indicate that changes in the use of intensity as a cue to fricative identification come about immediately when the need arises.

Table 1. *The number of participants in each group (number of males), and age and hearing information for each group.*

	N	**Age**	**Hearing (dB HL)**	
	(Male)	*Mean (SD)*	*Mean (SD)*	*range*
Younger, clean	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, clean	24 (8)	72.2 (4.6)	22.9 (8.8)	10.0 – 36.7
Older, filtered	24 (8)	72.2 (4.0)	23.1 (6.6)	13.3 – 36.7

2. Experimental set-up

2.1. 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, see Table 1. One group of older adults was newly recruited to participate in the filtered condition, while the older listeners in the clean condition were subsampled from [5], such that the number of older participants in the clean condition and filtered condition was equal and their age and hearing characteristics were matched as far as possible. Hearing sensitivity for all participants was assessed with a Maico ST20 portable audiometer (air conduction thresholds only) for octave frequencies from 250 Hz through 8 kHz. Mean pure-tone averages (averaged over participants' thresholds at 1, 2, and 4 kHz in their better ear) for the four participant groups are listed in Table 1. The two younger adult groups did not differ in their hearing sensitivity ($t(47) < 1$, $p > .1$) nor did the two older adult groups ($t(43) < 1$, $p > .1$). No participants wore hearing aids. Individual pure-tone average (over 1, 2, and 4 kHz in the better ear) of the participants was entered as an index of hearing loss in some of the analyses of the older listener results.

2.2. Materials

The materials were taken from [5]. There were four minimal pairs of /f/-final and /s/-final words: *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.

To investigate the role of intensity on the identification of final fricatives, we created versions of the test items of the four minimal pairs in which the intensity of the noise of the final fricative was varied. For each test item, the final natural fricative was excised, and the intensity of the word onset (e.g., *brie*) was set at 70 dB using Praat [9]. 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 naturally occurring intensities of the /f/'s and /s/'s in the eight stimuli. Subsequently, the intensity-modulated final fricatives were concatenated as final sounds to their corresponding /f/-final or /s/-final source word. The resulting stimuli were natural words, but with a final fricative that was intensity-modulated.

In Dutch, formant frequencies and transitions from the preceding vowel into a following /f/ or /s/ are fairly similar, with only a small difference in slope of F3 around 2200 Hz (a rise into a following /s/ and flat into a following /f/; [4]). To investigate the use of intensity in the absence of high frequency fricative noise, we need to keep this distinguishing formant transition available. Therefore, all stimuli are low-pass filtered at 2500 Hz using Praat. Note that, since the spectral information of /f/ is distributed uniformly over the spectrum, a difference in spectral information between the low-pass filtered /f/ and /s/ remains present in the signal.

2.3. Procedure

In all experiments, the participants were tested individually in a sound-treated booth. The stimuli were presented binaurally over closed headphones at a fixed maximum level of 70 dB SPL. 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 on the button box corresponding to the word they heard as fast and accurately as possible. They were not informed about the presence of intensity-modulated sounds.

The seven intensity-modulated versions 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 the two blocks. The stimulus lists in the low-pass filtered condition were identical to those in the clean condition, but using the low-pass filtered versions instead of the clean versions of the stimulus.

3. Results

All analyses were carried out using generalised linear mixed-effect models (e.g., [10]), containing fixed and random effects, using the logit link function and dummy coding. For each analysis, a best-fitting model is built using the fixed and random variables. Only statistically significant effects are part of the final statistical model and reported.

Figure 1 shows the proportion of /s/ responses for the seven intensity steps of the 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: the responses to the /s/-final stimuli are indicated with 's' for the younger and with 'S' for the older listener groups; the responses to the /f/-final stimuli are indicated with 'f' for the younger and with 'F' for the older listener groups. The results for the clean condition are indicated with solid lines and for the low-pass filtered condition with dotted lines. Table 2 displays the parameter estimates in the best-fitting models of performance for the four listener groups and two listening conditions. Because we took a subsample of the older listener group in the clean condition from [5], the results presented here differ slightly from those reported in [5]. Table 2 also presents the results for the age group comparison (younger vs. older adults, clean condition).

The responses to the intensity-modulated stimuli were subsequently analysed (the dependent variable is whether the response is /f/, coded as '0', or /s/, coded as '1'). The fixed predictors were Word (/f/-final word vs. /s/-final word; the latter is on the intercept) and Intensity level (step 1 (= 44 dB) through 7 (= 56 dB); a continuous variable with step 4 on the intercept). Subject and Minimal pair were the random factors. Age group (younger vs. older listener group; the former are on the intercept) was an additional fixed predictor in the age group comparison. Hearing loss (centralised to its mean) is an additional fixed predictor for the analysis of the results of the older listeners (there was little variation in hearing sensitivity among the younger participants, so hearing loss was not included in the analyses of the younger listeners).

3.1. Younger vs. older listeners, clean condition

Figure 1 clearly shows, not surprisingly, that there are significantly fewer /s/ responses to /f/-final words than to /s/-final words in the clean condition (see also Table 2, model A, Word). Important to our research question: younger and older listeners use intensity as a cue for fricative categorisation differently. The lack of a simple effect of Intensity level in model A indicates that younger listeners do not use intensity information for fricative categorisation, while the interaction Intensity level × Age group shows that older listeners do. The older listeners gave more /s/ responses to the /s/-final words with increasing intensity levels, but fewer /s/ responses to /f/-final words (the 3-way interaction). Older listeners thus rely on intensity as a cue to the identification of /s/ and not for the identification of /f/.

3.2. Younger and older listeners, clean condition

The difference in use of intensity as a cue for fricative categorisation was further investigated by analysing the results by the younger and older listeners separately. The analyses for the younger and older listener groups were similar (Table 2, models B and C), the only difference was that Hearing loss was included as a predictor for the older listeners but not for the younger listeners (as explained above).

As was found in the age group comparison, younger listeners (see model B) gave fewer /s/ responses to /f/-final words than to /s/-final words. Moreover, indeed, there was no simple effect of Intensity level. The older listeners (see model C) showed a similar result to the younger listeners for Word and Intensity level. There was a simple effect of Word, i.e., there are significantly fewer /s/ responses to /f/-final words than to /s/-final words, while there was no simple effect of Intensity.

[5] reported an effect of hearing loss. This result was still found with the smaller subsample of the older listeners: with increasing hearing loss, there were significantly fewer /s/ responses to the /s/-final words, but more incorrect /s/ responses to /f/-final words (Word × Hearing loss).

The three-way interaction (Intensity × Word × Hearing loss) indicates that intensity is used as a cue to fricative categorisation by listeners with hearing loss but differently for the /s/-final and /f/-final words. There is a marginal effect of Intensity × Hearing loss (p = .050, but p < .001 when including all 38 older listeners from [5]): with increasing hearing loss, there are more /s/ responses to /s/-final words with increasing intensity, i.e., listeners with hearing loss use intensity as a cue to identify /s/, with a higher intensity leading to relatively more correct /s/ responses. However, intensity is not used as a cue to identification of /f/ as is shown by the 3-way interaction: with increasing hearing loss, there are fewer /s/ responses to /f/-final words with increasing intensity. These intensity-related findings concur with the intensity-related findings in the age group comparison, suggesting that at least part of the differences in cue use found between the younger and older listeners are driven by the older listeners with hearing loss.

In short, younger listeners and older listeners with normal hearing do not rely on intensity as a cue to fricative identification, while older listeners with hearing loss do. The latter are influenced by intensity as a cue for the identification of /s/ and are not influenced by intensity differences for the identification of /f/.

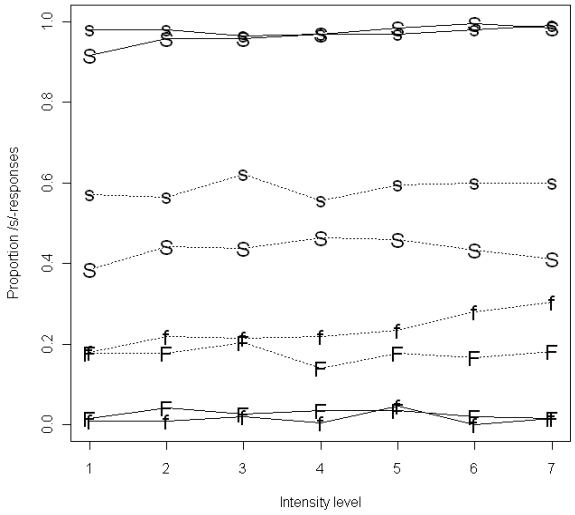

Figure 1. *Proportion of /s/ responses for the /s/- and /f/-final words for the older listeners (S, F) and the younger listeners (s, f) in the clean (solid lines) and filtered conditions (dotted lines).*

Table 2. *Fixed effect estimates for the best-fitting models of performance (number of observations). Dependent variable is the probability of /s/ responses (logit).*

Fixed effect	β	SE	$p<$
Model A: younger vs. older adult data, clean condition (5376)			
Intercept	3.887	.258	.001
Word (/f/-final word)	-8.222	.309	.001
Intensity	.028	.046	n.s.
Age group	-.004	.341	n.s.
Intensity × Word	-.003	.075	n.s.
Intensity × Age group	.162	.065	.05
Word × Age group	0.542	.416	n.s.
Intensity × Word × Age group	-.208	.098	.05
Model B: younger adult data, clean condition (2688)			
Intercept	3.758	.199	.001
Word (/f/-final word)	-7.976	.291	.001
Model C: older adult data, clean condition (2688)			
Intercept	4.281	.349	.001
Word (/f/-final word)	-8.506	.382	.001
Hearing loss	-.110	.030	.001
Intensity	.107	.066	n.s.
Intensity × Hearing loss	.012	.006	.1
Intensity × Word	-.084	.092	n.s.
Word × Hearing loss	.228	.037	.001
Intensity × Word × Hearing loss	-.019	.009	.05
Model D: younger adult data, low-pass filtered condition (2800)			
Intercept	.369	.324	n.s.
Intensity	.012	.015	n.s.
Word (/f/-final word)	-1.857	-.095	.001
Intensity × Word	.047	.023	.05
Model E: older adult data, low-pass filtered condition (2688)			
Intercept	-.733	.580	n.s.
Word (/f/-final word)	-2.013	.122	.001
Hearing loss	-.018	.071	n.s.
Word × Hearing loss	-.096	.019	.001

3.3. Younger and older listeners, filtered condition

When high-frequency information is removed from the speech signal, fricative categorisation becomes much harder. This is clearly illustrated in Figure 1 where the number of correct responses to the /s/-final and /f/-final stimuli drops dramatically (compare dotted vs. solid lines) for both the younger and the older listener groups, and this is particularly the case for the number of /s/ responses to /s/-final words.

Like was found for the younger listeners in the clean condition, there is no simple effect for Intensity level for the younger listeners in the filtered condition (Table 2, model D), showing that intensity is not used for the identification of /s/-final words. Importantly, however, there is an interaction Intensity × Word: younger listeners do rely on intensity as a cue for the categorisation of /f/, i.e., they give more /s/ responses to the /f/-final words with increasing intensity (see also the upwards turning dotted /f/-line for the higher intensities for the younger listeners in Figure 1). Younger listeners can thus be pushed to use intensity as a cue to fricative identity by removing high-frequency spectral information. However, unlike what was found for the older adults with hearing loss, who used intensity for the /s/-final words, intensity use in the younger adults leads to recognition errors for the /f/-final words. Thus, the younger adults were led to perceive the "loud f" fricative as /s/.

This raises the question how older listeners would perform when high-frequency information was removed (see Table 2, model E). Like was found for the older adults in the clean condition, fewer /s/ responses are given to /f/-final words. But regarding intensity, older listeners in the filtered condition no longer show effects of intensity modulation. Hearing loss did not affect identification of /s/-final words, but it did affect identification of /f/-final words: the interaction Word × Hearing loss shows that with increasing hearing loss, fewer (incorrect) /s/ responses are given to /f/-final words, so the removal of high-frequency information leads to more correct /f/ responses to /f/-final words for listeners with hearing loss. This is in contrast to the clean condition where increasing hearing loss resulted in an increase in incorrect /s/ responses to /f/-final words (see model C, Word × Hearing loss interaction).

4. General discussion

In a previous study [5], it was found that older listeners with hearing loss rely more on intensity information for fricative categorisation than older listeners with normal hearing. The current study addressed the question whether this cue use changes gradually due to progressive age-related hearing loss or whether the use of intensity for fricative identification may come about immediately when the need arises. This question was investigated by examining younger listeners' use of intensity as a cue to fricative identification and comparing it to a group of older listeners. Subsequently, younger and older listeners were tested on fricative identification when high-frequency spectral information was removed from the speech signal, which allowed us to investigate the reliance of younger and older listeners on high frequency spectral information in the speech signal for the categorisation of /f/ vs. /s/.

Phonetic categorisation experiments were conducted in which participants had to indicate whether they heard the /f/- or /s/-interpretation of four minimal Dutch word pairs. The intensity of the critical (natural) final fricative was either increased or decreased compared to its original intensity.

The results in the clean condition show that younger and older listeners with normal hearing do not use intensity as a cue to fricative identification. Like was reported in [5], older listeners with hearing loss, however, are influenced by intensity as a cue for the identification of /s/ but are not influenced by intensity differences for the identification of /f/.

In the low-pass filtered condition, as is clearly illustrated in Figure 1, the task of fricative categorisation becomes much harder, particularly for /s/. High-frequency information is thus highly important for the correct identification of /s/ but less so for /f/, as one would expect on the basis of the spectra for /s/ and /f/.

Low-pass filtering of the stimuli had a different effect on the two age groups regarding the use of intensity. Whereas in the clean condition, younger listeners did not use intensity as a cue, in the filtered condition, younger listeners did, they gave more /s/ responses to /f/-final words with increasing intensity, which led to recognition errors for the /f/-final words. Older listeners, on the other hand, did not use intensity as a cue to fricative categorisation in the filtered condition, not even the listeners with hearing loss, while these listeners did so in the clean condition.

Note again that age-related high-frequency hearing loss is by no means simulated by the removal of all high-frequency information as is done in the low-pass filtered condition. This could be an explanation why younger listeners use intensity differently from the older listeners with hearing loss in the clean condition. However, the removal of high-frequency spectral information only led to an increased reliance on intensity for the younger adults. We cannot exclude the possibility that older normal-hearing listeners might have shown an effect of intensity if the low-pass filtered condition had properly simulated age-related high-frequency hearing loss. Nevertheless, it is clearly not the case that older listeners immediately change their cue use when the need arises. Our results thus seem to suggest that, for older adults, intensity information may only gain in importance for fricative identification with longer exposure to a degraded speech signal.

5. Conclusion

Younger listeners and older listeners with normal hearing do not rely on intensity as a cue to fricative categorisation. Older listeners with hearing loss, however, are influenced by intensity as a cue for the identification of /s/ and not for the identification of /f/. Younger listeners can be pushed to use intensity as a cue for fricative categorisation when high-frequency information is removed from the speech signal, but the filtering manipulation did not have this effect on the older adults. The use of intensity is thus rather limited across groups and conditions. These results may suggest that the use of intensity as a cue to fricative categorisation shown by the older adults with hearing loss is only acquired gradually.

6. Acknowledgements

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

7. References

[1] Hedrick, M.S., Younger, M.S. “Labeling of /s/ and /ʃ/ by listeners with normal and impaired hearing, revisited”, J Speech Lang Hear Res. 46:636-648, 2003.

[2] Boothroyd, A. “Auditory perception of speech contrasts by subjects with sensorineural hearing loss”, J Speech Hear Res. 27:134-144, 1984.

[3] Dubno, J., Dirks, D., Langhofer, L. “Evaluation of hearing-impaired listeners using a nonsense-syllable test: II. Syllable recognition and consonant confusion patterns”, J Speech Hear Res 25:141-148, 1982.

[4] Rietveld, A.C.M., van Heuven, V.J. “Algemene Fonetiek”, Bussum: Dick Coutinho, 1997.

[5] Scharenborg, O., Janse, E. “Hearing loss and the use of acoustic cues in phonetic categorisation of fricatives”, Interspeech, Portland, OR, 2012.

[6] 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):1276-1284, 2002.

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

[8] Gordon-Salant, S., Frisina, R. D., Popper, A., & Fay, D. (Eds.). “The aging auditory system: Perceptual characterization and neural bases of presbycusis”, Berlin, Germany: Springer, 2010.

[9] Boersma, P., Weenink, D., “Praat. Doing phonetics by computer (Version 5.1)”, 2005.

[10] Baayen, R.H., Davidson, D.J., Bates, D.M. “Mixed-effects modeling with crossed random effects for subjects and items”, J Memory and Language, 59, 390-412, 2008.