EFFECTS OF AGE AND HEARING LOSS ON ARTICULATORY PRECISION FOR SIBILANTS

Xaver Koch[a,b] & Esther Janse[a,c]

[a]Radboud University, Nijmegen, The Netherlands; Centre for Language Studies
[b]International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands
[c]Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen
x.koch@let.ru.nl, e.janse@let.ru.nl

ABSTRACT

This study investigates the effects of adult age and speaker abilities on articulatory precision for sibilant productions. Normal-hearing young adults with better sibilant discrimination have been shown to produce greater spectral sibilant contrasts. As reduced auditory feedback may gradually impact on feedforward commands, we investigate whether articulatory precision as indexed by spectral mean for [s] and [ʃ] decreases with age, and more particularly with age-related hearing loss. Younger, middle-aged and older adults read aloud words starting with the sibilants [s] or [ʃ]. Possible effects of cognitive, perceptual, linguistic and sociolinguistic background variables on the sibilants' acoustics were also investigated.

Sibilant contrasts were less pronounced for male than female speakers. Most importantly, for the fricative [s], the spectral mean was modulated by individual high-frequency hearing loss, but not age. These results underscore that even mild hearing loss already affects articulatory precision.

Keywords: perception-production link, individual differences, sibilant, hearing loss, aging

1. INTRODUCTION

Adult aging may lead to several changes in speech, such as spectral modifications, altered voice characteristics and decreased speech rate [16]. Generally, imprecise articulation is a prominent perceptual feature of older adults' speech [7]. Auditory sensory decline has been argued to be a possible cause of the changes in their speech [3], besides cognitive and anatomical changes.

The finding that post-lingually deafened adults produce less pronounced consonant contrasts than healthy controls [11] emphasizes the role of auditory feedback for precise articulation. Further important evidence that perceptual differences are linked to production differences comes from Perkell and colleagues [13], showing that participants with good sibilant discrimination abilities also produce greater spectral sibilant contrasts. Thus far, investigations of age effects on speech production have mainly focussed on speech rate, fundamental frequency, vowel formant values, and voice onset time. However, sibilant fricatives, due to their spectral prominence in high-frequency ranges, can be expected to be the first to be impacted by age-related high-frequency hearing loss. Furthermore, sibilants are acquired relatively late in child language development and are often affected by speech disorders such as dysarthria or apraxia of speech.

We assume that the combination of the sibilants' complex articulatory movements and their dependence on precise auditory feedback relate to their vulnerability to disorders and to their relatively late acquisition age. A negative relation between sequence in language development and language decline has been shown for language impairments in dementia of the Alzheimer type [12], a neurodegenerative disease which has been linked to aging. Even healthy aging may be accompanied by reduced motor control that would be apparent particularly for sounds that are relatively difficult to produce, and that require high-frequency auditory feedback information.

The present study therefore investigates whether and how sibilant production may change across the adult life span (Research Question 1). Additionally, we investigate which individual cognitive, perceptual, linguistic and sociolinguistic speaker characteristics predict articulation precision (Research Question 2).

2. SPEAKERS

Three age groups were included (107 participants in total): 38 older adults (M^{age} = 67.1 yrs., *SD* = 4.7, 22 female), 34 middle-aged adults (M^{age} = 49.9, *SD* = 7.6, 21 female), and 35 younger adults (M^{age} = 21.4, *SD* = 2.6, 22 female). None of the participants wore hearing aids although six of them (one middle-aged and five older adults) met the Dutch hearing-aid criterion (pure-tone average over 1, 2 and 4 kHz $\geq$ 35 dB HL in either ear).

The speakers were sampled from a participant pool. All lived in the Nijmegen area, but came from different Dutch regions. Participants were asked to fill out a questionnaire on their language background and regional dialect. Participants were also asked to specify whether they spoke a Dutch dialect in everyday life or not (regionality self-rating).

3. PROCEDURE

3.1. Materials and speech recordings

Participants read ten monosyllabic target words (nine nouns, one adjective) embedded in a carrier phrase (*Ik zei ___ tegen hem* 'I said to ___ him'). The two target sounds [s, ʃ] appeared in five vocalic contexts (c.f. Table 1). Each target word was repeated five times. All stimulus pairs were near minimal pairs with the exception of one true minimal pair (*sop* vs. *shop*). Recordings were made in a sound-attenuated booth using a Samson QV head-set microphone and an Edirol R09 recorder (44.1 kHz sampling frequency, 16 bit resolution). Fifty filler sentences (ten nouns without sibilants in word-initial position, each repeated five times) were interspersed with the target sentences on a single pseudorandomized list. Sentences from this list were presented to participants one by one on a computer screen in a self-paced manner.

Table 1: Target words.

[s]			**[ʃ]**		
Saab	[saːp]	*car brand*	sjaal	[ʃaːl]	'scarf'
set	[sɛt]	'set'	chef	[ʃɛf]	'boss'
Sieb	[sip]	*name*	chic	[ʃik]	'modish'
sop	[sɔp]	'soap'	shop	[ʃɔp]	'shop'
soep	[sup]	'soup'	Sjoerd	[ʃuʁt]	*name*

3.2. Speaker abilities

Whereas there is a wealth of studies on individual predictors for speech comprehension, very little research has looked into relationships between speaker abilities and speech output [6]. We explore whether auditory, cognitive and linguistic abilities are associated with articulatory precision. The following five tests were administered:

1. *Pure-tone audiometry* to index hearing thresholds:
   · hearing level in decibel
2. *Digit Symbol Substitution Test* [17] performance to index processing speed:
   · number of correctly recoded symbols (within 2 min., out of 133)
3. *Vocabulary subpart of the Groningen Intelligence Test* [10] to index linguistic ability:
   · number of correct synonym answers (out of 20)
4. *Digit Span Test* [17] with backward recall to index working memory (visually administered):
   · percentage of correctly recalled items (12 items)
5. *Raven's Standard Progressive Matrices Test* [14] to index general non-verbal intelligence:
   · number of correct items (in 10 min., out of 60)

Table 2: Means and standard deviations of speaker abilities per age group.

Speaker abilities	Young adults *M (SD)*	Middle-aged adults *M (SD)*	Older adults *M (SD)*
2 kHz hearing threshold	7.14 (5.72)	10.29 (7.97)	20.39 (10.03)
4 kHz hearing threshold	4.43 (7.35)	17.65 (12.20)	29.74 (15.68)
6 kHz hearing threshold	10.57 (9.38)	20.88 (15.50)	34.89 (17.99)
8 kHz hearing threshold	6.43 (8.54)	25.15 (13.84)	45.53 (18.19)
Processing speed	87.26 (13.46)	76.15 (15.21)	64.05 (13.42)
Linguistic ability	13.83 (2.04)	15.76 (1.56)	16.68 (2.00)
Working memory	55.95 (18.81)	62.50 (24.38)	50.00 (18.78)
Non-verbal intelligence	44.54 (5.60)	38.82 (6.00)	31.58 (8.06)

3.3. Processing of acoustic data

The recorded sentences (107 speakers × 5 repetitions × 10 sentences = 5350) were pre-annotated using the automatic speech recognition plugin `praatalign` [9]. Sibilant and target word boundaries were checked and corrected (if necessary) manually for all productions but the forced alignment procedure was quite accurate (maximally 10-15 ms deviation from hand-annotated boundaries). Target word durations were extracted to model potential speech rate effects on articulatory precision.

We followed the procedures described by Forrest and colleagues [4] to derive spectral moments from the sibilant signals, except for the fact that frequencies were not transformed to Bark scale [8]. The middle third of each respective sibilant section was chosen as analysis interval to minimize coarticulation effects on the measurements. All analyses were done using `Praat` [2]. A pre-emphasis of 6 dB/octave for the frequencies above 80 Hz was applied to the analysis intervals. The resulting spectra were cepstrally smoothed (500 Hz) and spectral moments (in Hz) were calculated. Only the values of the first spectral moment (Center of gravity: henceforth, COG) of the sibilant productions were analyzed.

4. RESULTS

Statistical regression models were run using linear mixed-effect models in the program `R` with the `lme4` package [1] for the dependent variable spectral mean (COG). We started from a model containing interactions between sibilant identity ([s] vs. [ʃ]) and all experiment related control variables (vocalic context (± round), repetition number (1:5), trial position in experiment (1:100), speech rate in syllables per second) and gradually simplified this model using a backwards stepwise model selection approach. The modeling procedure was based on likelihood ratio tests to evaluate which interactions (sibilant identity × control variable), or control variables could be taken out without significant loss of model fit. The optimal random effect structure consisted of random intercepts for participants and items as well as random slopes for speech rate (correlated with random intercept for participants) and sibilant identity (no correlation with random intercept for participants).

Firstly, 22 target phrase productions with hesitations or slips of the tongue were excluded from the analyses. Secondly, spectral mean values above 10 kHz were excluded. Subsequently, outliers were removed separately for [s] and [ʃ] productions: COG values higher than 2.5 *SD*s above the respective means were excluded. Analyses are based on a dataset containing 2656 [s] and 2550 [ʃ] productions.

The resulting basic model (not reported here in detail as all effects are replicated in later models) showed significant effects of sibilant identity, vocalic context and trial position plus an interaction of sibilant identity × vocalic context (± round):

i. *sibilant identity effect:* higher spectral mean values for [s] compared to [ʃ] productions

ii. *vocalic context effect:* lower spectral mean values for the sibilants in +round (+back) vocalic context compared to −round context

iii. *trial position effect:* higher spectral means for trials later in the experiment

iv. *sibilant × vocalic context interaction effect:* stronger anticipatory coarticulation effects for [s] sibilants than for [ʃ] sibilants.

4.1. Modeling age and gender effects for sibilant productions

To test for basic speaker information effects on articulatory precision (Research Question 1), we added chronological age and gender of the speakers in one step to the basic model described above (simple effects, interactions with control variables and sibilant identity). Gender was included as a control variable as sibilant productions are known to differ between male and female speakers [5, 15]. The most parsimonious model resulting from adding age and gender effects is presented in Table 3.

Table 3: Model testing for age and gender effects in sibilant productions.

Fixed effects	β	*SE*	$p <$
Intercept	7125.26	105.11	
Sibilant identity	-1569.16	107.40	.001***
Vocalic context	-950.59	65.51	.001***
Trial position	0.66	0.27	.016*
Gender	-812.00	152.31	.004**
Gender × voc. context	240.19	39.85	.038*
Sibilant identity × gender	716.51	145.39	.001***
Sibilant identity× voc. context	708.07	96.76	.001***
Sibilant identity × voc. context × gender	-361.55	55.36	.001***

Reference levels: sibilant identity: [s], vocalic context:−round, gender: female; *P*-values were calculated using the `Anova` function of the `car` package (Type II Wald χ^2 test). Significance level notation: $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Age did not affect the sibilants' spectral mean, nor did it interact with any of the other predictor variables. However, adding participants' gender to the sibilant production model significantly improved the data fitting. In line with earlier studies [5, 15], male speakers showed lower spectral means for [s] productions than female speakers. Consequently, the acoustic contrast between [s] and [ʃ] productions is smaller for male than for female participants (cf. the sibilant identity × gender interaction). Men also show smaller coarticulation effects for [s] productions in the +round vocalic context compared to female speakers.

4.2. Modeling speaker ability effects on sibilant production

To investigate the role of speaker characteristics on sibilant production beyond age and gender effects, we carried out a third series of model comparisons. We did not model age and effects of speaker abilities simultaneously because hearing, processing speed, non-verbal intelligence and vocabulary size were all considerably correlated with age (Spearman's rank-order correlation tests: $|r| > .50$, $p < .001$). Thus, all background variables were added to the previous model (excl. age), as well as their interactions with sibilant identity. The resulting model (not shown here) showed that participants who categorized themselves as dialect speakers produced significantly lower spectral means than non-dialect participants (β = -305.44, SE = 121.96, $p < .05$). However, the absence of an interaction between regionality self-rating with sibilant identity implies that dialect speakers do not show reduced acoustic sibilant contrasts but rather shift both sibilants' acoustic spaces to lower frequencies.

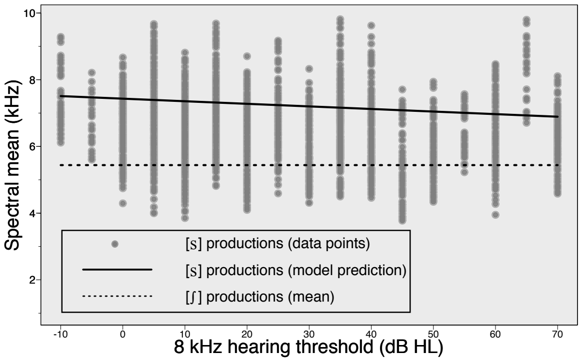
Interactions with sibilant identity were found for the two continuous predictors processing speed and 8 kHz hearing loss, whereas no simple effects were observed for these predictors. On the basis of these two interactions and after visual inspection of the relationship between hearing and the two sibilants' spectral means, we decided to run separate analyses for each sibilant to further investigate effects of speaker abilities on production of the sibilants.

Table 4: Model testing for speaker ability effects on [s] productions.

Fixed effects	β	SE	$p <$
Intercept	7509.06	134.49	
Gender	-834.02	144.00	.001***
Vocalic context	-961.62	57.10	.001***
Regionality self-rating	-373.57	146.57	.011*
8 kHz hearing threshold	-7.72	3.24	.018*
Gender × voc. context	242.38	43.86	.001***

Reference levels: vocalic context: −round, gender: female, regionality self-rating: no dialect use; P-values were calculated using the `Anova` function of the `car` package (Type II Wald χ^2 test). Significance level notation: ***$p < .001$,**$p < .01$, *$p < .05$.

Figure 1: Effect of high-frequency hearing loss on the spectral mean (COG) of [s] productions. Experimental data (n=2656) and model prediction shown, mean COG value for [ʃ] productions as reference.



Our analysis of the [ʃ]-production data did not substantiate the effects of processing speed, hearing loss or any other speaker ability measure on spectral mean. For the [s] productions only high-frequency hearing loss was a predictor of spectral mean frequency (and not processing speed or any other speaker ability measure): The higher participants' 8 kHz hearing threshold, the lower their spectral mean for [s]. This effect of high-frequency hearing loss ($\approx$ 8 Hz decrease in spectral mean per loss of 1 dB HL at 8 kHz) on the spectral properties of the [s] productions is illustrated in Figure 1. Table 4 shows the most parsimonious model on speaker abilities for the [s] productions (random effects structure: random intercepts for participants and items as well as a random slope for speech rate which was correlated with random intercept for participants).

To rule out that the hearing loss effect was solely due to the six participants who met the Dutch hearing-aid criterion, we also ran the above model on a dataset excluding these speakers. The effect of hearing acuity at 8 kHz on the spectral mean (COG) of the [s] productions was replicated in this subset.

5. DISCUSSION

Numerous studies have investigated sibilant production addressing different questions (e.g., on speech production modeling [13], or on sociophonetic variation [5, 15]). The present study was set up to investigate if changes over the adult life span influence articulation precision and to evaluate effects of individual speaker abilities on sibilant articulation. A standard sentence production paradigm was employed to elicit word-initial sibilant productions [s, ʃ] from a large sample of participants ($n > 100$), ranging in age between 18 and 78 years.

First, effects of vocalic context and speaker gender as found in other studies were replicated here, but the hypothesized age effect on sibilant articulation was not found (Research Question 1). Moreover, our data showed a gender by sibilant interaction effect, suggesting that the sibilant contrast was more pronounced for female than male speakers. Concerning our second research question on effects of speaker abilities on sibilant articulation, we found that high-frequency hearing loss modulated [s] productions. Thus, the sharpness of a speaker's [s] relates to the speaker's hearing acuity. Individual hearing acuity influences the auditory (feedback) information available from hearing one's own speech and from hearing other speakers. As we cannot be certain that the observed hearing acuity differences among speakers of our sample were actually acquired at an older age, our data indicate that high-fidelity auditory feedback is needed to either acquire or maintain precise articulation. Earlier research had shown effects of profound hearing loss on sibilant production. Our results, however, indicate that even mild (high-frequency) hearing loss modulates target production, particularly for targets with their distinct information in high-frequency spectral regions.

6. REFERENCES

[1] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. `lme4`: Linear mixed-effect models using Eigen and S4 (`R` package version 1.1-7), `http://CRAN.R-project.org/package=lme4`.

[2] Boersma, P., Weenink, D. 2014. `Praat`: Doing phonetics by computer [computer program, version 5.4.04], `http://www.praat.org/`.

[3] Cruickshanks, K., Wiley, T., Tweed, T., Klein, B., Klein, R., Mares-Perlman, J., Nondahl, D. 1998. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The epidemiology of hearing loss study. *J. Epidemiol.* 148(9), 879–886.

[4] Forrest, K., Weismer, G., Milenkovic, P., Dougall, R. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *J. Acoust. Soc. Am.* 84(1), 115–123.

[5] Fuchs, S., Toda, M. 2010. Do differences in male versus female /s/ reflect biological or sociophonetic factors? In: Fuchs, S., Toda, M., Zygis, M. (eds), *Turbulent sounds. An interdisciplinary guide*. Berlin: Mouton de Gruyter. 281-302.

[6] Haley, K., Seelinger, E., Mandulak, K., Zajac, D. 2010. Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *J. Phon.* 38(4), 548–554.

[7] Hartman, D., Danhauer, J. 1976. Perceptual features of speech for males in four perceived age decades. *J. Acoust. Soc. Am.* 59(3), 713–715.

[8] Jongman, A., Wayland, R., Wong, S. 1994. Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am.* 108(3), 1252–1263.

[9] Lubbers, M., Torreira, F. 2013-2014. `Praatalign`: An interactive `Praat` plug-in for performing phonetic forced alignment (version 0.9), `https://github.com/dopefishh/praatalign`.

[10] Luteijn, F., van der Ploeg, F. 1983. *Handleiding Groninger Intelligentietest (Manual Groningen Intelligence Test)*. Lisse, The Netherlands: Swets and Zeitlinger.

[11] Matthies, L., Sirsky, M., Lane, H., Perkell, J. 1994. A preliminary study of the effects of cochlear implants on the production of sibilants. *J. Acoust. Soc. Am.* 96(3), 1367–1373.

[12] Olga, V., Emery, B. 2000. Language impairment in dementia of the Alzheimer type: A hierarchical decline? *Int. J. Psych. Med.* 30(2), 145–164.

[13] Perkell, J., Matthies, M., Tiede, M., Lane, H., Zandipour, M., Marrone, N., Stockmann, E., Guenther, F. 2004. The distinctness of speakers' /s/– /ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *J. Speech Lang. Hear. Res.* 47, 1259–1269.

[14] Raven, J., Raven, J., Court, J. 2003. *The Manual for the Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview.* Harcourt Assessment San Antonio.

[15] Stuart-Smith, J. 2007. Empirical evidence for gendered speech production: /s/ in Glaswegian. In: Cole, J., Hualde, JI. (eds), *Laboratory Phonology. Vol 9*. Berlin: Mouton de Gruyter, 65-86.

[16] Torre, P., Barlow, J. A. 2009. Age-related changes in acoustic characteristics of adult speech. *J. Commun. Disord.* 42, 324–333.

[17] Wechsler, D. 2004. *Wechsler Adult Intelligence Test*. Harcourt Assessment (Lisse): Swets and Zeitlinger Amsterdam.