

A perceptual study of speaker age

Susanne Schötz

Dept. of Linguistics and Phonetics, Lund University
 Susanne.Schotz@ling.lu.se

Abstract

The present paper examines listeners' ability to judge speaker age from stimuli consisting of phonated isolated words, and also suggests subdivision (categorisation) of voices into typical and atypical groups, depending on the listeners' age estimating abilities.

1 Introduction

1.1 Purpose and hypotheses

The present paper, based upon my B.A. essay in phonetics (Schötz 2001) examines listeners' ability to judge speaker age from stimuli consisting of isolated words pronounced by elderly and young male and female speakers. Two hypotheses were tested:

1. Listeners are able to judge speaker age within ± 10 years of chronological age.
2. Some voices are more easily age-estimated than others, perhaps due to individual levels of instability and variability in the age-related cues of the voice, hence suggesting a subdivision into typical and atypical voices, depending on the degree of correspondence between perceived and chronological age.

This perceptual study serves as a preparatory investigation to an acoustical study of the same speech material, which is currently (April-May 2001) in progress.

1.2 The aging of the voice

Researchers have tried to relate certain acoustic and perceptual parameters of voice to the aging process. Although data are still sparse, literature often considers the following to be the main correlates of the aging voice: (a) fundamental frequency – pitch, (b) vocal intensity – loudness, (c) jitter and shimmer – roughness, (d) formant frequencies and spectral slope – voice quality, and (e) duration and pausation – speech rhythm and timing.

1.3 Previous studies

Most researchers agree that age-related changes in laryngeal and supralaryngeal structures affect the perceived acoustic properties of speech, that listeners' age judgements depend on both laryngeal and supralaryngeal cues (Jacques & Rastatter 1990), and that untrained listeners are able to judge speaker age within ± 5 years of chronological age at levels considerably better than chance (Hollien 1987). In most previous studies the listeners were asked only to place speakers in two (Jacques & Rastatter 1990, Ptacek & Sander 1966), three (Linville 1987) or four (Traunmüller & van Bezooijen 1994) age categories, rather than estimate age more precisely. Also, in most cases, the investigations focused on age perception of phonated (sometimes also filtered and whispered) vowels and/or whole sentences, and not isolated words. The phonated stimuli usually produced the best results, indicating that fundamental frequency is the most dominant cue for age estimation. Results from several studies have led researchers to suggest that there is a connection between speaker age

and listener age (Hartman 1979, Hollien & Tolhurst 1987, Huntley et al. 1980, Kreiman & Papcun 1985). Researchers have also reported several differences, although sometimes opposing, between listener groups (Hartman 1979, Jacques & Rastatter 1990, Linville & Korabic 1990).

2 Method

2.1 Speech material

The examined speech material consisted of the three Swedish words 'tack' 'thanks', 'rasa' 'fall down' and 'tusendollarsedlar' 'thousand-dollar bills', pronounced in the same dialect (småländska) by eight non-pathological speakers, giving a total of 24 words. All speakers (four female, age 73, 61, 30 and 28 years, and four male, age 71, 66, 27 and 21 years) were taken from the research project SweDia-2000's database. The speech material underwent a preliminary auditive analysis, which subdivided the speakers into four typical and four atypical voices in accordance with the second hypothesis of the study.

2.2 Procedure

In a perception test presented as a web page, 38 listeners (19 male, 19 female, age 14-60 years) were asked to age-estimate the 24 words, presented in a random order, by selecting one of 18 age alternatives, ranging from 10 to 95 years, from a menu. The results for each of the 24 words, for each of the eight speakers, and for speaker and listener groups, were analysed using calculations of the mean values, standard deviations (SD) and ranges, as well as the number and percentages of correct age estimates.

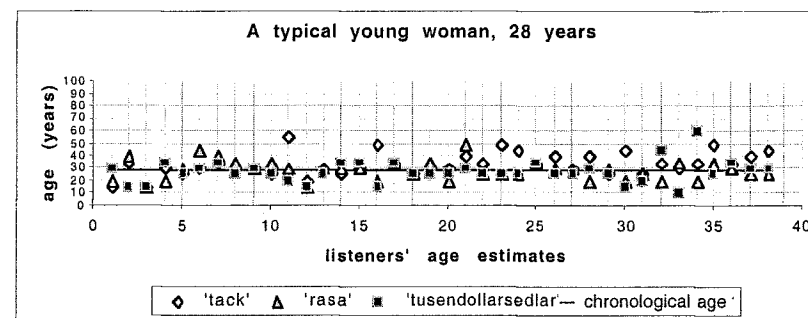


Figure 1. 38 listeners' age estimates of a typical young female speaker.

Young woman (28 years) (38 listeners)	'tack'		'rasa'		'tusendollarsedlar'	
	number	percent	number	percent	number	percent
Correct	9	24%	11	29%	6	16%
Correct ± 5 years	23	61%	25	66%	27	71%
Correct ± 10 years	28	74%	35	92%	30	79%
Mean	33,68		28,55		27,24	
Standard Deviation	9,42		7,96		9,13	
Max	55		50		60	
Min	15		15		10	
Range	40		35		50	

Table 1. 38 listeners' age estimates of a typical young female speaker.

3 Results and analysis

3.1 Typical and atypical voices

One of the main purposes of this study was to find out whether typical and atypical voices exist. The results indicate that such a subdivision actually may exist, and that the preliminary subcategorization of the speakers turned out to be accurate enough to support the hypothesis. The four typical speakers did get a greater amount of correct age estimates as well as lower SD and range values than the four atypical speakers. Figures 1-2 and Tables 1-2 illustrate age estimates for two of the speakers of the study: a typical young woman and an atypical elderly man.

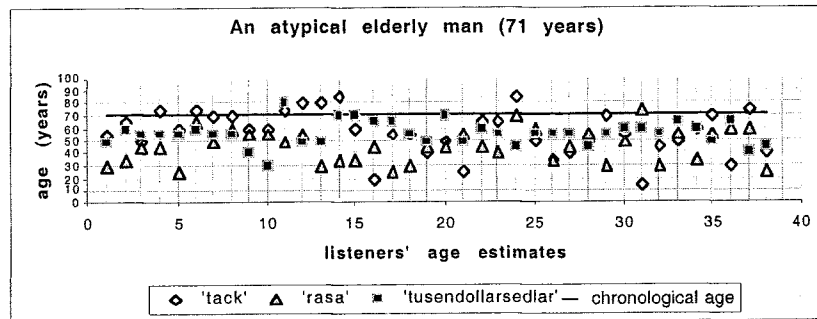


Figure 2. 38 listeners' age estimates of an atypical elderly male speaker.

Elderly man (71 years) (38 listeners)	'tack'		'rasa'		'tusendollarsedlar'	
	number	percent	number	percent	number	percent
Correct	4	11%	1	3%	3	8%
Correct ± 5 years	10	26%	3	8%	8	21%
Correct ± 10 years	18	47%	7	18%	14	37%
Mean	56,97		45,79		55,66	
Standard Deviation	17,61		13,18		9,53	
Max	85		75		80	
Min	15		25		30	
Range	70		50		50	

Table 2. 38 listeners' age estimates of an atypical elderly male speaker.

3.2 Different sets of comparisons between speaker and listener groups

Differences between results from different speaker groups (in terms of the relationship between chronological and estimated age) and listener groups (regarding their age estimation abilities), such as male vs. female or young vs. older, were found in several of the previous studies. Other researchers, however, found no such indications (Kukul et al. 1980). This investigation did only find minor differences, which is probably due to the small size of the study. Also, a comparison of groups of trained vs. untrained listeners was made, but no notable differences were found.

4 Conclusions and discussion

Listeners were not always able to judge speaker age within ± 10 years of chronological age, but the results show that for some speakers (i.e. the typical speakers) the estimates were considerably better than chance, which in this case could be considered as $3/18 \approx 16,7\%$ as

there were 18 alternatives in the perception test. Between 19 (50%) and 35 (92%) of the 38 listeners made correct age estimates within ± 10 years for the four typical speakers, whereas the result obtained for the atypical voices was poorer; only between 7 (18%) and 22 (58%) correct age estimates within ± 10 years.

The better results for the typical speakers suggest that a subdivision of voices into typical and atypical categories could help in further understanding the cues of age perception, for instance when integrating age-related parameters in speech synthesis.

The monosyllabic word 'tack' got the poorest results for all speakers. It did not seem to hold enough information about the speaker to enable several of the listeners to make reasonably good age estimates. This indicates that more than a monosyllabic word is probably needed in order to judge age correctly.

The fact that the results of different studies do not always agree, indicates that more research in this field is needed to fully understand the acoustic and perceptual age-related properties of speech.

References

- Decoster, Wivine. 1998. *'Akoestische kenmerken van de ouder wordende stem'*. Leuven: Leuven University Press. (summary in English).
- Hartman, D. 1979. 'The perceptual identity and characteristics of aging in normal male adult speakers'. *Journal of Communication Disorders* Vol 12, 53-61.
- Hollien, H. & Tolhurst, G. 1987. The aging voice. In: Weinberg, B, (ed.), *Transcripts of the seventh symposium on care of the professional voice, part II: life span changes in the human voice*. New York: The Voice Foundation. 67-73.
- Hollien, Harry. 1987. "'Old Voices": What do we really know about them?' *Journal of Voice* Vol. 1, No 1, 2-17.
- Huntley, Ruth, Hollien, Harry & Shipp, Thomas. 1980. 'Influences of listener characteristics on perceived age estimations'. *Journal of Voice* Vol. 1, No 1, 49-52.
- Jacques, Richard D & Rastatter, Michael P. 1990. 'Recognition of speaker age from selected acoustic features as perceived by normal young and older listeners'. *Folia Phoniatrica* Vol 42, 118-124.
- Kreiman, J. & Papcun, G. 1985. 'Voice discrimination by two listener populations'. Paper presented at ASA, Austin, TX.
- Kukul R, Hutchinson, D. & Duane, M. 'Listener group comparisons of perceived speaker age'. Paper presented at the ASHA convention in Detroit, Mi.
- Lindblad, Per. 1992. *Rösten*. Lund: Studentlitteratur.
- Lindblad, Per. 1998. *Talets akustik och perception*. Department of Linguistics and Phonetics. Göteborg University.
- Linville, SE. & Fisher, H. 1985. 'Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females'. *J. Acoust. Soc. Am.* Vol 78, 40-48.
- Linville, Sue Ellen. 1987. 'Acoustic - Perceptual Studies of Aging Voice in Women'. *Journal of Voice* Vol. 1, No 1, 44-48.
- Ptacek, Paul H. & Sander, Eric K. 1966. 'Age Recognition from Voice'. *Journal of Speech and Hearing Research*. Vol. 9, 273-277.
- Schötz, Susanne. 2001. 'Röstens ålder - en perceptionsstudie'. B.A. essay in Phonetics. Department of Linguistics and Phonetics. Lund University.
- Trautmüller, Hartmut. 1997. 'Perception of speaker sex, age, and vocal effort'. <http://www.ling.su.se/staff/hartmut/F97.pdf>.
- Trautmüller, H. & van Bezooijen, R. 1994. 'The auditory perception of children's age and sex'. In *Proceedings ICSLP-94, vol. 3, The Acoustical Society of Japan*: 1171 - 1174.