Published in final edited form as:

Logoped Phoniatr Vocol. 2016 October ; 41(3): 101–105. doi:10.3109/14015439.2015.1009160.

Listener Estimations of Talker Age: A Meta-Analysis of the Literature

Eric J. Hunter,

Department of Communicative Sciences and Disorders, Michigan State University, East Lansing MI

Sarah Hargus Ferguson, and

Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, UT

Catherine Anne Newman

Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, UT

Abstract

Numerous studies, most of them cross-sectional studies using one sample per talker, have demonstrated that listeners make relatively accurate age judgments from hearing talkers' voices. The current study analyzed the results of these direct age estimation studies to better characterize the perception of talker age over a larger number of individuals. A review of the direct age estimation literature was performed. Datasets from seven papers were reconstituted and an analysis of the combined data (meta-analysis) including 530 data points was conducted. The reconstituted and combined data included talkers aged 10–90. Listeners appeared to overestimate age when talkers were younger and to underestimate it when talkers were older.

Keywords

Age estimation; Meta-analysis

INTRODUCTION

Beginning in the fifth decade of the human lifespan, specific degenerative physiological changes (e.g., muscle atrophy, nerve and cell dystrophy, edema) affect speech. Such deterioration includes dysphonia (1); slower speech and elongated words (2); tremor, instability, and aspiration (3); reduced loudness (4); and voice fundamental frequency changes (5). Furthermore, decreased pulmonary function may cause weak voice and shortened phrase lengths (6).

Correspondence concerning this article should be addressed to: Eric Hunter, Department of Communicative Sciences and Disorders, Michigan State University, 1026 Red Cedar Road, Room 113, Oyer Speech & Hearing Building, East Lansing, MI 48824, Contact: ejhunter@msu.edu.

DECLARATION OF INTEREST

The authors report no declarations of interest outside the following stated research funding.

These age-related physiological changes are encoded acoustically in the speech signal and are somehow perceived as aging in talkers. To investigate both the accuracy with which listeners can perceive age as well as what those encoded acoustic characteristics might be, there have been a range of age identification studies. Such studies have included a variety of talker characteristics (e.g., gender, age range) and listener characteristics (e.g., gender, age, number), as well as a variety of speech materials (e.g., sustained vowels, spontaneous speech) and perceptual tasks (e.g., category judgments, direct age estimates). Despite the numerous methodological differences among these studies, categorical judgments of age have clearly been more accurate for connected speech materials than for sustained vowels. However, it is important to note that accuracy also seems to decrease as the number of age categories increases.

With few exceptions (7), most category judgment studies have shown that listeners tend to underestimate the age of older talkers. For example, Caruso, Mueller, and Xue (8) found that two older talkers were much more likely to be judged as young or middle-aged than to be judged as old. This age underestimation for older talkers is also evident in direct age estimation studies, most of which have covered several decades of talker ages (9–12).

Notwithstanding the importance of any individual study, the results of these studies have not yet been combined into a concise knowledge base. This type of consolidation would not only make the data more usable, but would also highlight similarities and differences in methodologies. In the current study, a literature review was conducted of direct age estimation studies of English talkers; data were then extracted from reviewed papers and concatenated into a single data set. The results can be used to increase our understanding of age estimation and perception of talker age over a large number of speakers. Eventually, such studies can be used to illuminate which particular vocal characteristics listeners use to identify age changes.

METHODS

A systematic review of age estimation studies and reports using native English-speaking talkers and listeners was conducted using the PubMed database during July and August 2012. Key words that were used, either in combination or singly, included: "age and perception," "perception of voice," "aging voice," and "perceptions of aging voice." Using those key words and a few well-known, seminal articles, the Google Scholar "cited by" feature was used to find articles citing those identified by PubMed. The reference lists from each of the papers were also examined to identify any additional papers. Thus, it should be noted that the goal of the review was not to conduct a traditional, repeatable literature survey, because such a survey would only identify studies specifically related to the keywords. Rather, our goal was to identify as many potential studies as possible.

Articles with pertinent data, whether in tabular or figure form, were noted for extraction. Before being compiled into a table, data were captured from digital scans of the articles using Snagit (13). Next, the digital capture was verified using visual inspection. Scatter plot data were captured and recreated using a MATLAB script called *GRABIT* (14), which, after loading digital images, allows users to calibrate axis settings and zoom or scroll in the image

while logging locations (in terms of the calibrated x/y axes) of highlighted data points. From these logged data locations, tables were recreated from the scatterplot data. Because of the ability to zoom into the scanned image while getting the data, the click location error was negligible. However, if scatterplots contained overlapping data points, best judgments were made by the research technician (and reviewed by another person) regarding the precise location of these points.

For each of the papers found and datasets extracted, data were analyzed both individually and collectively. Additionally, the collected dataset was segmented into five-year increments (per *chronological age* of the talkers.). For each dataset and increment, the average estimated *perceived age* was calculated. Furthermore, the average difference between perceived age and chronological age was calculated, as were the average absolute difference and the correlation coefficient between perceived age and chronological age.

RESULTS

Upon completion of the literature search, 23 peer-reviewed articles (published in English, with native English-speaking speakers and listeners) spanning from 1966 to 2010 were selected. Each article was read and categorized according to stimuli, listeners, and perceptual testing procedures. Thirteen of the 23 articles included direct age estimates in the experimental procedures. Of these 13 papers, 7 displayed the resulting data in a table or scatter plot rather than presenting only descriptive statistics. Table 1 is a summary of the seven papers, including number of talkers and age.

Listener age range varied across the studies. Table 2 summarizes the number and gender of listeners in each study. Although most studies used students or “young adults” as listeners, Huntley et al. (12) used a range of listeners aged 9 to 84 years while Eppey and Mueller (15) used two sets of listeners (25 aged 18 to 22 years and 18 aged 61 to 84 years) to assess potential listener age bias effects on age estimations. Hartman (11) did not state the age range of his listeners, though ten were speech-language pathologists. Speech samples to be judged were presented in random order for all studies (except for Moran, McCloskey, & Cady, (16), which did not state how materials were presented). Stimulus presentation was carried out via earphones in two studies and via loudspeakers in another four; for the remaining study the means of presentation was not stated. Only two articles gave any indication of the level at which the speech samples were presented: Huntley et al. (12) used a presentation level of 70–75 dBA, while Hartman (11) presented materials at a “comfortable level.” Stimuli included steady vowels (15), 30-s from 5-minute recordings of spontaneous speech (11), and samples excerpted from read passages. Excerpted read speech samples consisted of either one sentence (9,12), two sentences (10,17), or whole paragraphs (15,16). All of the studies in Table 2 included running speech of some sort.

A meta-analysis of these seven published articles was accomplished. A total of 530 data points (perceived age to chronological age comparisons) were conglomerated; this total matched the number of talkers from Table 1. The correlation coefficient between perceived age and chronological age was 0.85. Figure 1 depicts the chronological age and the

perceived age estimates for all data points with a straight line representing where the chronological age and the perceived age would be equivalent.

The combined dataset was then divided into five-year talker chronological age increments from 10 to 90 years. Overall, each half-decade contained an average of 31 talkers. Between the ages of 20 and 85, the lowest number of talkers per half-decade was 18. The average estimated perceived age for each half-decade are depicted in Figure 1 as large diamonds. Overall, the conglomerated dataset shows that listeners overestimated age when talkers were young and underestimated age when talkers were older. When talkers were between the ages of 45 and 55 years, the averaged perceived age and chronological age are nearly the same.

The average chronological age of the talkers in the seven papers was 54.63 years, while the average perceived age was 51.86. The signed difference score (perceived age – chronological age) was calculated for each of the 530 data point; these are plotted in Figure 2, with solid circles for every five years of chronological age. When considering talkers over a wide range of ages, both age overestimation for young talkers (~5 to 10 years) and age underestimation for older talkers (~8 to –12 years) are apparent. Overall, the age estimation differences for younger and older talkers mostly cancel each other, resulting in an overall average underestimation of age of -2.77 SD 9.56 years. However, the signed difference hides the absolute difference for each age estimate. Absolute differences ($|\text{perceived age} - \text{chronological age}|$) were therefore calculated for all data points; these are shown in Figure 2 as open circles. The average absolute difference was 7.98 SD 5.94 years. The most accurate half-decades occurred between 35 and 55 years, with an absolute age estimation difference of about 6 years during this chronological age span, while the least accurate half-decades occurred in the oldest talkers (85–90 y/o) with an absolute difference of more than 11 years. The average signed difference per half-decade decreased almost linearly as talkers aged ($r^2=0.94$), decreasing by 1 year of difference every 3.5 years of chronological age. On the other hand, the relationship between average absolute differences and chronological age had a parabolic shape, with a 3rd-order polynomial showing a much better fit to the data ($r^2=0.86$) than a linear function ($r^2=0.64$).

DISCUSSION

The current study reviewed and concatenated previous studies that examined whether listeners are capable of making accurate judgments of a talker's age from hearing his or her voice, permitting a detailed look at listeners' absolute age-estimation abilities over the human lifespan of talkers. The age estimate data show a shift from age overestimation for younger talkers to age underestimation for older talkers. When age estimates are compared to actual chronological ages, the data plot of signed differences (diamonds in Figure 1) features a crossover point at which estimated and perceived age are equal for the average group of listeners. This crossover point occurs at approximately 50 years. When individual listener responses are compared in terms of the absolute difference between perceived and chronological age (open circles in Figure 2) rather than the signed difference, however, a slightly different perspective on the accuracy of age estimation emerges. Although signed differences suggest that listeners make the most accurate age judgments when talkers are between 45 and 70 years old, the absolute differences reveal that individual listeners still

misjudge talker age by 5 to 10 years in this age range. That is, the apparently high level of accuracy for this span of ages is misleading when absolute differences are considered.

The present meta-analysis combines results from many previous studies. Variability exists among the previous studies' results in part because of the methodological variability among numerous independent investigators. It should also be noted that most of the listeners in the various studies were young adults/students. The meta-analysis thus primarily reflects young ears, making it possible that the results do not apply to older listeners. However, when Eppley and Mueller (15) compared young and elderly listeners, they found that the two groups of listeners underestimated the age of older talkers to the same degree. Another caveat involves our technique for reconstituting the published data; it is possible that small errors may have occurred in spite of the steps taken to avoid them. However, we anticipate that any such errors would be small and have little effect on the overall trends observed here. Finally, it is possible that gender would have an effect on age estimation. However, there was not significant enough information from the papers to adequately attribute the estimated age to any gender effect.

If considered in relation to data on age-related vocal changes, the results of our meta-analysis could lead to the identification of the specific vocal characteristics that best reveal the age of individual talkers. Knowing these characteristics could be useful in a variety of situations. For example, a health care provider interacting with and assessing the health of a patient might note a discrepancy between how old the patient sounds and how old the patient actually is. This discrepancy might then be recognized as a symptom of an underlying health condition. In addition, the current results could guide the development of speech signal processing algorithms in which computers are used to estimate age or synthesize aged voices.

Acknowledgments

We would like to acknowledge Catherine Mellum and Russell Banks for their help with extracting points from old figures.

Research reported in this publication was in part supported by the National Institute On Deafness And Other Communication Disorders of the National Institutes of Health under Award Number R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCE LIST

1. Golub JS, Chen P-H, Otto KJ, Hapner E, Johns MM. Prevalence of Perceived Dysphonia in a Geriatric Population. *J Am Geriatr Soc.* 2006 Nov 1; 54(11):1736–1739. [PubMed: 17087701]
2. Ramig LA. Effects of physiological aging on speaking and reading rates. *J Commun Disord.* 1983 May; 16(3):217–226. [PubMed: 6875037]
3. Ramig LA. Effects of physiological aging on vowel spectral noise. *J Gerontol.* 1983 Mar; 38(2): 223–225. [PubMed: 6827040]
4. Baker KK, Ramig LO, Sapir S, Luschei ES, Smith ME. Control of vocal loudness in young and old adults. *J Speech Lang Hear Res.* 2001; 44(2):297. [PubMed: 11324652]
5. Stathopoulos ET, Huber JE, Sussman JE. Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age. *J Speech Lang Hear Res JSLHR.* 2011 Aug; 54(4):1011–1021. [PubMed: 21173391]

6. Huber JE, Spruill J. Age-Related Changes to Speech Breathing With Increased Vocal Loudness. *J Speech Lang Hear Res.* 2008 May 27; 51(3):651–668. [PubMed: 18506042]
7. Ptacek PH, Sander EK. Age recognition from voice. *J Speech Hear Res.* 1966 Jun; 9(2):273–277. [PubMed: 5925532]
8. Caruso AJ, Mueller PB, Xue A. Relative contributions of voice and articulation to listener judgements of age and gender: preliminary data and implications. *Voice.* 1994; 3:1–9.
9. Shipp T, Hollien H. Perception of the aging male voice. *J Speech Hear Res.* 1969 Dec; 12(4):703–710. [PubMed: 5385720]
10. Ryan EB, Capadano HL. Age perceptions and evaluative reactions toward adult speakers. *J Gerontol.* 1978 Jan; 33(1):98–102. [PubMed: 618973]
11. Hartman DE. The perceptual identity and characteristics of aging in normal male adult speakers. *J Commun Disord.* 1979 Feb; 12(1):53–61. [PubMed: 422747]
12. Huntley R, Hollien H, Shipp T. Influences of listener characteristics on perceived age estimations. *J Voice.* 1987; 1(1):49–52.
13. SnagIt. TechSmith Corporation, Version 10.0. 2011
14. Doke J. GRABIT. Matlab Central File Exchange. 2007
15. Eppley BD, Mueller PB. Chronological age judgments of elderly speakers: the effects of listeners' age. *Contemp Issues Commun Sci Disord.* 2001; 28:5–8.
16. Moran MJ, McCloskey L, Cady B. Listener Age Estimates of Elderly African American and Caucasian Male Speakers. *J Cross-Cult Psychol.* 1995 Nov 1; 26(6):751–758.
17. Ryan WJ, Burk KW. Perceptual and acoustic correlates of aging in the speech of males. *J Commun Disord.* 1974 Jun; 7(2):181–192. [PubMed: 4841332]

Figure 1.

Conglomerated data from seven studies yielding more than 600 data points. The large diamonds represent the average response for each half decade. The solid line represents where perceived and chronological age would be the same.

Figure 2.

Accuracy of Direct Age Estimation per half decade presented in two ways. The filled circles represent the average signed difference between chronological age and estimated age; the open circles represent the average absolute difference between chronological age and estimated age. R2 values correspond to the linear fit lines; the polynomial fit is also provided for visualization purposes.

Table 1

Direct age estimation articles reconstituted from printed manuscripts.

Reference	# Talkers	Min CA	Max CA	Avg CA	Std CA	Range CA
Eppley & Mueller (2001)	18	71.0	82.0	76.4	3.8	11.0
Hartman (1979)	37	25.0	66.0	44.0	12.2	41.0
Huntley et al. (1987)	96	19.6	89.3	54.0	20.0	69.7
Moran et al. (1995)	56	62.0	86.0	73.4	6.2	24.0
Ryan & Burk (1974)	80	40.8	79.6	59.6	11.8	38.8
Ryan & Capadano (1978)	69	11.9	70.5	34.4	16.9	58.6
Shipp & Hollien (1969)	174	20.2	89.5	54.7	19.9	69.3

Listed are the number of talkers, the minimum chronological age (CA), maximum CA, average CA, standard deviation of CA, and the range of CA.

Table II

Listeners used in direct age estimation articles reconstituted from printed manuscripts.

Reference	# Listeners	Age Range	Gender	Stimuli Presented
Eppley & Mueller (2001)	43 (25 & 18)	18–22, 61–84	both	Vowel; complete GP
Hartman (1979)	30 (20 & 10)	not stated, SLPs	not stated	Spontaneous speech (30 sec)
Huntley et al. (1987)	120 (4 × 30)	9–15, 20–30, 40–50, 60–84	both	3 rd sentence of RP
Moran et al. (1995)	17	“students”	female	1 st paragraph of RP
Ryan & Burk (1974)	20	“young adult”	female	4 th and 5 th sentences of RP
Ryan & Capadano (1978)	98	“students”	both (85% male)	1 st two sentences of a read passage
Shipp & Hollien (1969)	30	“students”	not stated	3 rd sentence of RP

GP = Grandfather Passage. SLP = speech-language pathologist. RP = Rainbow passage.