The Impact of Visual and Auditory Cues in Age Estimation

Kajsa Amilon, Joost van de Weijer, Susanne Schötz

The impact of visual and auditory cues in age estimation

Kajsa Amilon, Joost van de Weijer, Susanne Schötz

Centre for Languages and Literature

Lund University, Sweden

The impact of visual and auditory cues in age estimation

When meeting a person for the first time, we form an opinion of him or her in just a few seconds (Gladwell, 2006). In this process, age estimation is an important part, which provides a large amount of information about the person we meet.

When making this estimation, visual cues (such as wrinkles, hair style, posture, and clothes) probably have a large impact. Recent studies suggest that people are able to determine a persons age by approximately five years, from looking at a picture of their face (Vestlund, 2004). Auditory cues, such as word choice and voice, also play a role, but yield less exact estimates than visual cues. Studies have shown that, when visual cues are not available, people can estimate a person's age with an absolute difference of 10 years. This is not as good as estimates based on visual cues, but still considerably better than chance (Ptacek & Sander, 1966).

Acoustic characteristics of voice contain information about a speaker's age. Physical changes during the ageing process (decreasing lung capacity, calcification of the larynx, loss of muscular control) influences our voices to various degrees. Audible changes include a decrease in vocal intensity, lowering of vocal pitch, and narrowing of the vocal pitch range. Furthermore, voices tend to become harsher. Finally, elderly people often speak slower than younger people (Schötz, 2001). However, Brückl and Sendlmeier (2003) found no difference in articulation rate between young and old women during read speech.

Recently, Schötz (2006) described how these changes in voice can be synthesised using data-driven formant synthesis. In her study, acoustic parameters were extracted from the Swedish word 'själen' (the soul), spoken by four differently aged females of the same family. The information was subsequently used to synthesize voices of different ages.

The accuracy with which people can estimate a person's age, either from auditory or from visual information, has been studied into great detail. The following effects have come forward.

The expert-effect

The expert-effect means that training makes perfect. Studies have shown that, for instance, professionals who work with specific age groups, are more accurate at estimating the age within that age range than others (Lindstedt, 2005). Vestlund (2006) for example, showed that sales persons at the Swedish alcohol retail company (who are not allowed to sell alcohol to persons under 20) were better than a non-expert group at estimating the ages of people between 15 and 24 years old. A similar finding is reported by Nagao and Kewely-Port (2005) who compared native speakers of Japanese with native speakers of American English. Both groups were asked to estimate the age of Japanese and of American English speakers. Overall, the subjects were more correct when estimating the age of a speaker of their own language. Furthermore, the Japanese listeners were better than the English listeners at estimating the ages. This might due to the fact that Japanese more frequently interact with elderly people (more than half of the Japanese subjects lived with at least one grandparent). In another study (Dehon & Bredart, 2001) with pictures instead of living persons, Caucasian subjects performed better at estimating the age of other Caucasians than they were on estimating the age of people of other races. A third variant of the expert effect is that it is easier to estimate

the age of persons your own age (George & Hole, 1995; Vestlund, 2004). In a line-up confrontation experiment where subjects were to recognize a culprit among other persons, the subjects were more often correct when pointing out a perpetrator their own age (Wright & Stroud, 2002).

Age

Overall, young adults tend to outperform older adults at age estimations, both from pictures (Vestlund, 2004) and sound (Linville, 1987). In a forensic context, young people have been found to be more reliable than elderly people as eyewitnesses (Wright & Stroud, 2002). Mulac and Giles (1996) added to the age effect by showing that someone's voice is not perceived as the person's chronological age, but how old the person feels. That is, if a person feels younger than his or her chronological age, this will be reflected in this person's voice and consequently, the listener will perceive him or her as younger.

Stimulus duration

The duration of the acoustic stimulus has an impact on how well subjects are able to estimate age from a speaker. Schötz (2005) found that the absolute difference between biological age and perceived age increases significantly from 6.5 to 9.7 years when the stimulus duration (spontaneous speech) decreases from 10 to 3 seconds. Nagao & Kewely-Port, (2005) also found that estimations improved when subjects received more contextual information.

Sex

Conflicting results exist as to whether it is easier to estimate a man's age or a woman's age. Dehon and Bredart (2001) and Krauss et al. (2001), on the one hand, found that the ages of

4

men were easier to estimate than the ages of women. Schötz (2005) and Vestlund (2006), on the other hand, found that subjects were better at estimating women's ages.

Dialect

Elderly speakers are more likely to have a strongly pronounced dialect than younger speakers, as well as men tend to be more dialectal than women (Stölten & Engstrand, 2003; Stölten, 2001). The strength of dialect consequently is an auditory cue when estimating the age of a speaker.

General effects

Several studies of both age estimation from pictures (Vestlund, 2004) and from acoustic cues (Schötz, 2005; Cerrato et al., 2000) suggest that the age of younger people often is overestimated, while the age of older people is underestimated. This indicates that subjects tend to place speakers in the middle range. Cerrato et al. (2000) received most correct estimations of people in the age group 46-52. Dehon and Bredart (2001), however, found that poor age estimations from faces are most often overestimations.

In all, age estimation is a fairly well examined research area. However, most studies only concern the impact of either voice or looks. It is very common, though, that both auditory and visual information are available at the same time. An unexplored question is whether the voice of a speaker has an impact on the estimation of a person's age when information about the person's looks is simultaneously available. Is it easier to make a correct estimation when having access to both sources of information? And if the voice does not agree with the looks, what affects the perceiver more, picture or sound?

Differences in estimations made from voices and pictures were examined by Krauss et al. (2002). They carried out two experiments in which they compared age estimation from photographs and from voices. In the first experiment, the subjects heard two sentences read by two different speakers. They were then shown two photographs which they had to match with the two voices. The subjects selected the correct photograph more than 75% of the time, which is reliably above chance level. In the second experiment a group of participants were exposed to the same sentences as in the first experiment but this time they were asked to estimate the speaker's age, height and weight. A comparison group made the same estimations from pictures of the speakers. The results showed that the estimations made from the photographs were only marginally better than the ones made from the voice samples.

It is clear that higher accuracy in age estimating is achieved when perceivers have access to a person's looks instead of only the voice (cf. an absolute difference of five respective ten years). It is unclear, however, how much simultaneously available visual and auditory cues contribute individually to the estimation of speaker age. The present study addresses this issue in a series of five related experiments. We recorded 14 differently-aged women on video tape. These video clips were subsequently presented to five groups of participants, every time in a different modality: 1) as still pictures, 2) as soundless videos, 3) as audio clips, 4) as the full video with the original sound, and, 5) as the full video but with the voice of the speaker replaced with the voice of one of the other speakers.

This design makes it possible to directly compare the estimation of age when a speaker is shown with or without voice. The goal of the Experiment 5 was to establish whether the participants' estimations could be altered by giving the speakers a voice that sounded either considerably older or younger (as was revealed by the results of Experiment 3) than their own

6

voices. The results of this experiment were then compared with those of Experiment 4 in order to establish whether the estimated age of a speaker had changed. If so, that would indicate that the voice contributes to age estimation. Experiment 2 was included in the present study mainly to establish whether the extra information provided by soundless video yielded significantly better estimates than still pictures.

Method

Speaker characteristics

Since previous research (Dehon & Bredart, 2001; Krauss et al., 2001; Schötz, 2005; Vestlund 2006) had shown a difference in the ability of estimating men and women, we decided on using only female speakers for the present study. Because of the pronounced change of voice in puberty only adults participated. The age interval of approximately five years was chosen because previous studies show that participants are able to estimate a speaker's age with an accuracy of approximately five years. The women's ages were: 17, 25, 31, 35, 38, 45, 47, 54, 55, 58, 65, 73, 76, 81, representing the entire adult age range. All women spoke Southern Swedish dialect. They were non smokers and did not have any known pathological voice disorder. Nor did any of them have a cold at the time of the recording. The speakers had not received any instructions about visual appearance (clothing, make up, hair style). They all wore casual clothing, and, according to us, they did not look or sound older or younger than their chronological age.

Equipment/recording conditions

The recordings were made at some time between 10 AM and 4 PM in an isolated recording booth. The speakers sat down on a chair in front of a neutral light background where they were recorded in portrait form. Recordings were made with a DV camera (Panasonic NVGS-180) and an external wireless microphone (Sony UTX-H1), mounted on a tripod. All women were first given the opportunity to look at the test sentence for as long as they needed. A sign with the text was placed next to the camera, so that the speakers could read it out loud and did not need to memorize the exact text. In order to facilitate dubbing we tried to obtain test sentences that were all read at approximately equal speed. The reading speed of the first

speaker (age 25) was used as a reference during the subsequent recordings. The other speakers were first asked to read in their own tempo, and then, if necessary, told to increase or decrease their speaking rate to end up as close as possible to the reference time, without their speaking style becoming unnatural.

Material

The sentence used for the investigation was "En undersökning som denna är viktig för förståelsen av folks röster i olika åldrar" (A study like this is important for the understanding of people's voices in different ages). In order to minimize the effect of dubbing a new voice onto a speaker, the sentence had few visibly salient speech sounds (for example, bilabial, labiodental). Furthermore, the sentence was not too long (a duration of about five seconds). According to previous research results (Schötz, 2005) five seconds is rather short to make a fair estimation of the age. In order to compensate for this, the participants in our study were allowed to watch or listen to the stimuli as many times as they wished.

Preparation of the materials

The video recordings were transferred to a Macintosh computer for preparation for the five experiments. Preparation was done using the video-editing program iMovie, and the speech-editing software Audacity. 14 clips with similar durations that seemed most suitable for the dubbing experiment were first selected from the available material. These clips were used for Experiments 2 and 4. For Experiment 1, a neutral-looking still picture of each speaker was taken from the original clips, and for Experiment 3, the mono sound signals were extracted and saved as audio files.

9

Our aim for Experiment 5 was to combine speakers with a new voice that sounded between 15 and 20 years younger or older than the speaker's own voice. It turned out that this was not possible for all speakers. In spite of the effort taken to equalize reading speed and utterance duration during the recordings, there was, inevitably, considerable variation in phrasing and pausing, which made it impossible to find an appropriate combination with a new voice for some speakers. We found 10 combinations of speakers and voices that appeared good enough for participants not to notice that the voice of the speaker was dubbed. In order to mask the fact that some video clips were dubbed, five dubbed clips were combined with five original video clips, and presented in two experimental sessions. The voices were dubbed onto the recordings with iMovie. In some cases, minor changes (e.g., shortening of fricative durations, shortening or lengthening of pauses) were made to the sound signal with Audacity so that the fit between sound and image improved.

Subjects

To participate in the sound only experiment (Experiment 3), the original video (Experiment 4) experiment or the dubbed video (Experiment 5) experiment, participants could only be native Swedish speakers. This was to prevent a possible 'other-language effect' affecting the results. In the other experiments, there were no nationality restrictions for participating. All participation was voluntary.

Procedure

Most participants were tested in a computer room at Lund University. An additional number were tested at other locations. In those cases, a portable Macintosh with headphones was used.

Participants received written instructions, which were similar for all five experiments, except for the description of the stimuli. The instruction for Experiment 5 (dubbed video) was identical to that of Experiment 4 (original video).

At the start of the experiments, all participants provided information about their age, sex, native language and dialect. They were not allowed to ask any questions during the experiment.

The video clips and still pictures were displayed on the computer screen. The picture was approximately 20x20 centimetres. Sound was presented via headphones at a comfortable listening level, but participants were allowed to adjust the volume if they wanted to. For every stimulus, the participants typed the speaker's estimated age in exact whole years (e.g., 24, 82, 57), and rated their judgment on a scale from 1 to 5, indicating whether they were very confident (5) about their answer or not at all (1). The stimuli were presented in random order. The results were saved automatically on the computer's hard disk.

At the end of the experiment, the subjects filled in a questionnaire with the following six questions: *What information did you use for your estimation? What was your main cue? Did you find anything hard in particular? Did you find any of the persons harder to estimate than the others? Did you react to anything in particular? Did you recognise any of the women in the clips?* Responses to speakers that the participant had recognised were excluded from further analysis. The answers to these questions were used for evaluating the results. At the end of Experiment 5, the participants were additionally asked if they had noticed that the videos were dubbed. If so, the participant was excluded from the analysis. All experiments took approximately 15 minutes to complete.

Analysis

Two main comparisons were carried out. The first comparison concerns the accuracy of the participants' estimates. As in several previous studies (cf. Vestlund, 2006; Schötz, 2005; Braun & Cerrato, 1999; Nagao & Kewley-Port, 2005), the error is calculated as the absolute difference between the speaker's chronological age and the perceived age. If this difference is not lower on average in Experiment 4 (original video) than in Experiment 2 (soundless video), then it seems that auditory cues are more or less neglected when visual cues are available. The second comparison concerns the difference in estimated age of the speakers in Experiment 4 (original video) and those in Experiment 5 (dubbed video). If the estimated age of a speaker does not change when she is given a voice that is considerably older or younger than her own voice, that would additionally suggest that auditory information is neglected when visual information is available.

In all, 141 persons participated. The majority were students or staff at the Centre for Languages and Literature at Lund University. None of them participated in more than one experiment. None of them reported any hearing disorder.

Experiment 1: Still pictures

14 women and 9 men participated in Experiment 1. Their average age was 33 years. 28 were Swedish, and 5 had a different nationality. In total there were 322 responses, of which 13 (4.0%) were excluded, either because the participant had recognised a speaker or because of equipment failure. The results (given in Table 1) show that there was an average absolute difference of 5.7 years. Furthermore, the average difference between chronological age and perceived age was –2.7, indicating that the participants somewhat underestimated the age of the speakers.

Table 1 here

Experiment 2: Soundless video

14 women and 14 men participated in experiment 2. Their average age was 27. 19 were Swedish, and 9 had a different nationality. 380 responses were analyzed. 12 responses were excluded (3.1%) because the participant had recognised a speaker or because of equipment failure. The overall results of Experiment 2 were similar to those of Experiment 1. As in Experiment 1, there was a tendency to underestimate the age of the speakers and the absolute error was somewhat over five years. The average confidence ratings were similar in both

experiments, suggesting that the participants did not find the videos more difficult to assess than the still pictures.

Table 2 here

Experiment 3: Sound only

16 women and 12 men participated in Experiment 3. Their average age was 29. There were 392 responses, of which 11 (2.8%) were excluded, either because the participant had recognised a speaker or because of equipment failure.

Table 3 here

The sound only experiment appeared, like previous research had shown, to constitute the most difficult assignment for the participants. With an average absolute difference of approximately ten years, the error was twice as large compared to the results of Experiments 1 and 2. This difficulty was nevertheless not reflected in particularly lower confidence ratings in Experiment 3 compared to those in Experiments 1 and 2. Remarkably, the tendency to underestimate the speakers' ages was even more pronounced in Experiment 3. The average estimated age was even further below the speakers' chronological age in Experiment 3. This was especially the case for the speakers aged 65 and 81, whose voices were estimated more than 20 years below their chronological ages.

Experiment 4: Original video

9 women and 12 men participated in this experiment. Their average age was 25. A total of 289 responses were analyzed. 5 responses (1.7%) were excluded either because the participant had recognised a speaker, or because of equipment failure. The results are given in Table 4.

Table 4 here

The average absolute difference between estimated age and chronological age was 5.1 years, which is comparable to the values found in Experiments 1 and 2. Interestingly, the two speakers aged 65 and 81 were not estimated considerably younger in Experiment 4 than in Experiment 1, in spite of the fact that their voices sounded more than 20 years younger in Experiment 3. This is a first indication that voice did not play a significant role in the judgment of age.

Comparison of the results of experiments 1 to 4

The absolute differences between chronological age and perceived age in Experiments 1 (soundless video), 2 (still picture) and 4 (original video) were tested in an ANOVA with 'experiment' as fixed factor, and 'participants' and 'speaker age' as random factors. Overall, the differences between the experiments were significant ($F[3, 52.048] = 5.308$, $p < .05$). Post-hoc comparisons using Tukey's HSD procedure showed that the differences between Experiments 1, 2, and 4 were not significant. Stated otherwise, soundless video did not yield more accurate responses than still pictures, and video with sound did not yield more accurate responses than video without sound.

The average confidence ratings, on the contrary, were not significantly different across the first four experiments ($F[3, 52.340] = 0.229$, $p > .05$).

The results of the first four experiments do not suggest that speakers' voices contributed substantially to the estimation of age when visual information was available. This suggestion is supported by the answers that the participants gave the questionnaires. Only four participants of Experiment 4 wrote that they also had paid attention to the voices of the speakers while making their judgments. Most participants, though, only mentioned visual cues, such as wrinkles, hair colour, hair style, and so on.

Experiment 5: Dubbed video

24 women and 17 men participated in the final experiment. They had an average age of 35.2. years. A total of 200 responses were analyzed. 5 responses were excluded because the participant had recognised the speaker, or because of equipment failure. Table 5 gives an overview of the results.

Table 5 here

The main comparison of interest is how the speaker's perceived age changes as a consequence of the new voice. This change is shown in the third and fourth column of the table. For all but two of the speakers (aged 38 and 55), the average perceived age in Experiment 5 was more than five years higher than the perceived age in Experiment 4. For the other speakers the differences in perceived age were smaller than five years.

Comparison of experiments 4 and 5

16

The estimated ages of the speakers in Experiment 4 were compared with those in Experiment 5 in order to establish whether the new voices had a significant effect on the participants' judgments. This turned out not to be the case. The differences in perceived age were not significant ($F[1,17.018] = 0.046$, $p > .05$). Overall, the confidence ratings were a bit lower in Experiment 5 than in Experiment 4. This difference was not significant either ($F[1, 16.987] = 1.602$, $p > .05$).

Discussion

The purpose of the present study was to investigate the impact of voice in age estimation. To what extent is age estimation influenced by voice when we see and hear a person simultaneously?

The results of our experiments are generally consistent with those of previous studies. Participants were more accurate at estimating the age of the speaker based on visual information compared to auditory information. Soundless video did not yield more accurate estimations than still pictures.

It was found that the accuracy with which participants estimated the ages of the fourteen speakers did not improve significantly when sound was added to the videos. Furthermore, deliberately manipulating the voices in order to make the speaker appear younger or older, did not have a significant impact on the participants' estimations. This was corroborated by the qualitative evaluations given by the participants at the end of the experiment. They reported that they paid attention to visual details, and hardly ever did they mention voice as an important characteristic. Taken together, these results suggest that voice does not contribute substantially to the estimation of age, and that estimations are based on visual appearance mainly, if not to say only.

This conclusion is in a sense limited, since voice is not the only auditory cue that contains information about speaker age. Other cues, such as speaking rate, word choice, dialect, were kept constant to a maximal extent in the present study. Nor do the present results show whether the same conclusion can be drawn for male speakers.

Even if voice does not appear to contribute significantly to the estimation of speaker age, it will be interesting to establish whether the impact of voice is larger for the estimation of other speaker characteristics, such as personality or mood. The present methodology of dubbing voices onto different speakers appears to be a promising way of investigating speaker judgments. Nevertheless, the number of dubbing possibilities was limited with the available material. In order to overcome these limitations, a tool such as the speaker age synthesizer (Schötz, 2006) can be of great value in the future for carrying out new experiments with more speakers and naturalistic material.

Reference List

Braun, A. and Cerrato, L. (1999). Estimating speaker age across languages. In Proceedings of ICPhS 99, pages 1369–1372. San Francisco, CA.

Brückl, M. & Sendlmeier, W. (2003). Aging female voices: An acoustic and perceptive analysis. Institut für Sprache und Kommunikation, Technische Universität, Berlin, Germany. Voqual '03, Geneva, August 27-29.

Cerrato, L., Falcone, M. & Paoloni, A. (2000) Subjective age estimation of telephonic voices. *Speech Communication*, 31: 107-112.

Dehon, H. & Bredart, S. (2001). An 'other-race' effect in age estimation from faces. *Perception*, 26 (9): 1107-1113.

George, P. & Hole G. (1995). Factors influencing the accuracy of age estimations of unfamiliar faces. *Perception*, 24 (9): 1059-1073.

Gladwell, M. (2006). *Blink: The Power of Thinking without Thinking*. New York: Time Warner Book Group.

Krauss, R., Freyberg, R. & Morsella, E. (2002). Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology*, 38: 618-625.

Lindstedt, R. (2005). Finns det experter på åldersbedömning? (Are there experts on the estimation of age?) Unpublished Paper. Department of Education and Psychology, University of Gävle.

Linville, S. (1987) Acoustic-perceptual studies of an aging voice in women. *Journal of Voice*, 1: 44-48.

Mulac, A. & Giles, H. (1996). You're only as old as you sound: Perceived vocal age and social meanings. *Health Communication*, 3: 199-215.

Nagao, K. & Kewley-Port, D. (2005). The effect of language familiarity on age perception. International Research Conference on Aging and Speech Communication, Bloomington, Indiana, October 10.

Ptacek, P. & Sander, E. (1966). Age recognition from voice. *Journal of Speech and Hearing Research,* 9: 273-277.

Schötz, S. (2001). Röstens ålder – en perceptionsstudie (The age of the voice – a perceptual study). Unpublished Paper. Department of Linguistics, Lund University.

Schötz, S. (2005). Effects of stimulus duration and type on perception of female and male speaker age. Proceedings, FONETIK, 2002. Department of Linguistics, Gothenburg University.

Schötz, S. (2006). Data-driven formant synthesis of speaker age. Working Papers 52, 105-108. Lund University.

Stölten, K. (2001). Dialektalitet som ålders- och könsmarkör i arjeplogsdialekten: ett auditivt test och mätningar på preaspiration och VOT. (Dialect as a marker of age and sex in the Arjeplog dialect: an auditory experiment and measurements of pre-aspiration and VOT.) Unpublished paper. Department of Linguistics Stockholm University.

Stölten, K. & Engstrand, O. (2003). Effects of perceived age on perceived dialect strength: A listening test using manipulations of speaking rate and F_0. *Phonum*, 9: 29-32.

Vestlund, J. (2004). Åldersbedömning av ansikten – precision och ålderseffekter. (Estimation of age from faces – Precision and age effects). Unpublished paper. Department of Education and Psychology, University of Gävle.

Vestlund, J. (2006). Åldersbedömning av ansikten – expertkunskaper, könseffekter och jämnårighetseffekter (Estimation of age from faces – expert effects, sex effects, and same-age effects). Unpublished paper. Department of Education and Psychology, University of Gävle.

Wright, D. & Stroud, J. (2002) Age differences in line-up identification accuracy: people are better with their own age. *Law and Human Behaviour*, 26: 641-654.

TABLE 1: Results of Experiment 1 (Still picture)

Chronological age	Perceived age	Difference	Absolute difference	Confidence rating
17	21.6	4.6	4.7	3.30
25	24.2	-0.8	2.5	3.18
31	25.9	-5.1	5.1	3.58
35	36.2	1.2	3.5	3.26
38	37.0	1.0	2.4	3.26
45	41.2	-3.8	4.6	3.23
47	47.6	0.6	4.6	3.17
54	43.8	-10.2	10.9	2.65
55	50.5	-4.5	5.4	3.23
58	50.9	-7.1	8.0	2.95
65	69.4	4.4	6.1	3.13
73	68.3	4.7	6.0	3.22
76	70.7	-5.3	7.7	3.05
81	74.0	-7.0	8.0	3.35
Total		**-2.7**	**5.7**	**3.18**

TABLE 2: Results of Experiment 2 (Soundless video)

Chronological age	Perceived age	Difference	Absolute difference	Confidence rating
17	20.0	3.0	3.0	3.29
25	24.5	-0.5	1.9	3.57
31	27.2	-3.8	4.6	3.08
35	33.1	-1.9	3.6	3.21
38	35.9	-2.1	4.1	3.13
45	40.8	-4.2	5.4	3.22
47	48.4	1.4	4.4	3.30
54	47.9	-6.1	7.1	2.93
55	51.0	-4.0	6.0	3.14
58	51.0	-7.0	8.0	2.81
65	67.4	2.4	5.3	2.89
73	67.6	-5.4	6.6	3.07
76	73.0	3.0	4.9	3.18
81	77.2	-3.8	5.6	3.14
Total		**-2.5**	**5.0**	**3.14**

TABLE 3: Results of Experiment 3 (Sound only)

Chronological age	Perceived age	Difference	Absolute difference	Confidence rating
17	18.5	1.5	3.4	3.64
25	24.0	-1.0	2.1	3.74
31	29.3	-1.7	5.6	3.11
35	35.2	0.2	7.0	2.96
38	26.3	-11.7	12.2	3.27
45	40.8	-4.2	6.6	2.86
47	42.0	-5.0	6.8	2.85
54	47.2	-6.8	8.9	3.11
55	46.2	-8.8	9.9	3.00
58	50.7	-7.3	10.0	2.67
65	42.9	-22.1	22.1	2.78
73	76.2	3.2	6.7	3.64
76	66.6	-9.4	10.6	3.18
81	55.7	-25.3	25.3	2.93
Total		**-7.0**	**9.7**	**3.13**

TABLE 4: Results of Experiment 4 (Original video)

Chronological age	Perceived age	Difference	Absolute difference	Confidence rating
17	18.3	1.3	2.0	3.86
25	24.6	-0.4	2.0	3.29
31	27.1	-3.9	4.3	3.20
35	32.7	-2.3	3.0	3.35
38	34.7	-3.3	5.0	3.05
45	39.9	-5.1	6.1	3.19
47	45.9	-1.1	2.9	3.00
54	47.0	-7.0	7.65	2.95
55	50.1	-4.9	5.1	3.48
58	51.4	-6.6	7.3	2.86
65	62.6	-2.4	4.0	3.20
73	70.0	-3.0	5.9	2.95
76	69.9	-6.1	8.3	3.14
81	74.0	-7.0	7.7	3.19
Total		**-3.7**	**5.1**	**3.19**

TABLE 5: Results of Experiment 5 (Dubbed video)

Chronological age	Combined with voice perceived as	Perceived age in original video	Perceived age in dubbed version	Confidence rating
25	40.8	24.6	26.6	3.33
35	50.7	32.7	34.8	3.14
38	24.0	34.7	33.1	3.35
38	47.2	34.7	40.3	2.95
45	29.3	39.9	40.1	3.05
47	29.3	45.9	48.4	3.25
55	76.2	50.1	58.1	2.70
58	66.6	51.4	55.0	2.70
65	66.6	62.6	65.4	3.19
76	50.7	69.9	70.2	2.75
Total				**3.04**