



Age estimation in foreign-accented speech by non-native speakers of English

Dan Jiao^{a,b,*}, Vicky Watson^{b,c}, Sidney Gig-Jan Wong^{b,c}, Ksenia Gnevsheva^d, Jessie S. Nixon^e

^a College of International Studies, Yangzhou University, Yangzhou 225009, China

^b Department of Linguistics, School of Language, Social and Political Science, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

^c New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

^d School of Literature, Languages and Linguistics, Australian National University, Canberra ACT 2601, Australia

^e Quantitative Linguistics Group, University of Tübingen, Wilhelmstr.19, 72074 Tübingen, Germany

ARTICLE INFO

Keywords:

Age estimation
Foreign-accented speech
Non-native listeners
Linguistic familiarity
Acoustic cues of age

ABSTRACT

Listeners are able to very approximately estimate speakers' ages, with a mean estimation error of around ten years. Interestingly, accuracy varies considerably, depending on a number of social aspects of both speaker and listener, including age, gender and native language or language variety. The present study considers the effects of four factors on age perception. It investigates whether there is a main effect of speakers' native language (Arabic, Korean and Mandarin) even when speaking a second language, English. It also investigates a particular speaker-listener relationship, namely the degree of linguistic familiarity. Linguistic familiarity was expected to be greater between Mandarin and Korean than between Mandarin or Korean and Arabic. In addition, it considers the effect of the acoustic cues of mean fundamental frequency (F0) and speech rate on age estimates. Fifteen Arabic-accented, fifteen Korean-accented and twenty Mandarin-accented English speakers participated as listeners. They heard audio stimuli produced by forty-eight speakers, equally distributed between native Arabic, Korean and Mandarin speakers, reading a short passage in English. Listeners were instructed to estimate speakers' ages in years. Listeners' age estimates and reaction times were recorded. Results indicate a significant main effect of speaker native language on perceived age such that Mandarin speakers were estimated to be younger than Arabic speakers. There was also a significant effect of linguistic familiarity on age estimation accuracy. Age estimates were more accurate with greater linguistic familiarity, i.e., native Korean and Mandarin listeners estimated ages of speakers of their own native languages more accurately than native Arabic speakers' ages and vice versa. In terms of acoustic cues, mean F0 and speech rate were significant predictors of age estimation. These effects suggest that in perception, age may be marked not only by biological changes that occur over the lifetime, but also by language-specific socio-cultural features.

1. Introduction

Listeners are generally capable of assigning an age category to unfamiliar speakers by listening to their voice (Cerrato et al., 2000; Shipp and Hollien, 1969). Previous studies on voice-age perception have found that native listeners are able to discriminate between younger and older voices with a high degree of accuracy (Cerrato et al., 2000; Hughes and Rhodes, 2010; Ptacek and Sander, 1966; Ryan and Burk, 1974; Shipp and Hollien, 1969). Ptacek and Sander (1966) found that accuracy in categorising between younger (under 35 years) and older (over 65 years) speakers was as high as 99% when listening to spoken words, although it was reduced to 78% when listening to prolonged vowels, rather than full words. However, when listeners estimate speakers' age to the year,

accuracy is substantially lower (Amilon et al., 2007; Hartman, 1979; Huntley et al., 1987; Krauss et al., 2002; Moyses et al., 2014; Schötz, 2005). Several researchers have found an absolute estimation error of around ten years (Amilon et al., 2007; Moyses et al., 2014).

A number of social factors may affect age estimation accuracy, including the speaker's age and gender and the listener's age and gender. The effect of speaker age on estimation accuracy appears to be non-linear. The age of younger adults tends to be overestimated, while the age of older adults tends to be underestimated (Braun, 1996; Cerrato et al., 2000; Harnsberger et al., 2010; Huntley et al., 1987; Schötz, 2005; Shipp and Hollien, 1969; Waller et al., 2015). This suggests a central tendency effect (Hollingworth, 1910). Listener age, on the other hand, seems to be negatively correlated with age estimation accuracy,

* Corresponding author at: College of International Studies, Yangzhou University, Yangzhou 225009, China.

E-mail addresses: dan.jiao@yzu.edu.cn (D. Jiao), vmw36@uclive.ac.nz (V. Watson), ksenia.gnevsheva@anu.edu.au (K. Gnevsheva), jessie.nixon@uni-tuebingen.de (J.S. Nixon).

<https://doi.org/10.1016/j.specom.2018.12.005>

Received 12 April 2018; Received in revised form 6 December 2018; Accepted 19 December 2018

Available online 20 December 2018

0167-6393/© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)

but also interacts with speaker age (Huntley et al., 1987; Linville and Korabic, 1986; Moysse et al., 2014). Estimating young speakers' age seems to be more challenging for older listeners (referred to as 'own-age bias'), while younger listeners are equally as accurate as older listeners at estimating older speakers' ages. There may also be an effect of speaker gender, with the age of female voices being better estimated than the age of male voices in older speakers (Hughes and Rhodes, 2010; Moysse, 2014). Speaker's smoking habits may also influence perceived age, in that smokers are estimated to be older than non-smokers of the same calendar age, probably because of the effect of smoking on acoustic characteristics of one's voice (Braun, 1996).

There are several acoustic correlates of speaker's age that may influence listeners' age estimations, including fundamental frequency (F0) and speech rate (Harnsberger et al., 2008, 2010; Hartman, 1979; Hartman and Danhauer, 1976; Linville, 1987; Ptacek and Sander, 1966; Schötz, 2007; Shipp et al., 1992; Torre and Barlow, 2009; Waller et al., 2015). F0 is found to decrease with age for females, but increase with age for males (Torre and Barlow, 2009). Speech rate appears to slow for both males and females as they age (Harnsberger et al., 2010). Faster speech rate is associated with lower age estimates especially in older speakers (Waller et al., 2015). In addition, voice quality cues such as tremor and noise are found to be significant predictors of perceived age (Harnsberger et al., 2010; Ryan and Burk, 1974). Most prior studies have investigated native speaker age estimation. So far, it is not yet known whether such acoustic cues as F0 and speech rate also affect age estimation of non-native speakers, and by non-native listeners.

1.1. Cross-linguistic effects in age perception

Relatively few studies have explored the effects of the listener and speaker's native language on age estimation accuracy. Braun and Cerato (1999) investigated native German and Italian listeners' age estimation of native German and Italian language speakers. There were no significant differences in age estimation accuracy when listening to their own language versus the other language. This null effect could be attributed to several factors. It may be due to there being substantial language familiarity between speakers of German and Italian, since Italy shares a border with both Austria and German-speaking Switzerland and has its own German-speaking region. Italian speakers may be relatively familiar with German speakers and vice versa. Additionally, it may be due to the fact that Italian and German are related Indo-European languages or due to cultural similarity, as Italy and Germany are neighbouring European communities with a substantial shared history, so that age cues do not vary substantially between the two language communities. Another possible explanation is that the quality of recordings was degraded, which may have had an effect on age perception accuracy. In the study, materials were either recorded over the telephone (Italian speakers) or first recorded to tape and then the recording was played over the telephone and rerecorded (German speakers). So although this study reported a null result, it does not provide strong evidence against a potential effect of language on age perception.

On the other hand, Nagao and Kewley-Port (2005) showed an effect of native language on estimation accuracy when native Japanese and native English participants listened to single vowels (/i/), trisyllabic non-words (/bisisi/) and full Japanese and English sentences produced by English and Japanese native speakers. The native Japanese listeners were more accurate overall than native English speakers. With both groups combined, listeners were significantly better at estimating their native language than the unfamiliar language. Nagao (2006) argues that language familiarity plays an important role in age perception and that listeners might rely on their language-specific perceptual experience when judging the age of a speaker whose linguistic background is unfamiliar to them.

1.2. Speaker accent effects

Two recent studies have investigated age estimation in foreign-accented versus native speech. Rodrigues and Nagao (2010) reported a difference in age estimation of foreign-accented versus native-accented English speech by native English listeners, and found an effect of listeners' foreign-accented English experience on their age perception. The spoken English stimuli were produced by five native Arabic speakers and five native English speakers (aged 18–79 years) selected from the Speech Accent Archive (Weinberger, 2015). Thirty native English speakers with differing degrees of experience with foreign-accented speech participated online via Amazon Mechanical Turk. The task not only involved age estimation, but also accent ratings and native language identification. The correlation between speakers' chronological age and estimated age was higher for native English listeners listening to English stimuli rather than the Arabic-accented stimuli, and was also higher for listeners with more experience of foreign-accented English, compared to the less experienced group. This study involved a small number of speakers over a wide age range (age range is specified only jointly, not separately for each language group). However, the results suggest the effects of language experience and foreign accent on accuracy of age estimation.

Further support for an effect of accent nativeness comes from a study of native English and Japanese listeners' age estimation of native British English and native Japanese speakers speaking English (Bürkle and Gnevsheva, 2017). This study examined listeners' age estimation accuracy and reaction times. The results showed that Japanese speakers were estimated to be younger and their reaction times were significantly slower for English listeners. The lowest correlation between real and estimated age was found for English listeners listening to Japanese speakers, followed by Japanese listeners listening to both English and Japanese speakers, and the highest correlation was for English listeners listening to English speakers. However, with only two speaker and listener groups, of which one was native English speakers, it is unclear whether this effect stemmed from linguistic familiarity or speaker nativeness.

1.3. The present study and hypotheses

To date, very few studies have investigated age estimation in foreign-accented speech, and those that have employed native listeners only. However, in modern society, second language speakers of English are as likely to communicate with other second language speakers as they are with native speakers of English, which makes the interaction between speaker and listener native languages an important question to investigate. In addition, most age estimation studies have focused on English or European languages or comparing between English and one Asian language. As far as we are aware, no previous study has explored cross-language effects in native speakers of different Asian languages. Neither has any previous study investigated both non-native speakers and listeners of different languages. Investigating more than two language backgrounds allows us to isolate the effect of linguistic familiarity as opposed to non-nativeness. Therefore, in the present study, second language speakers of English participate as both speakers and listeners: Native Arabic, Korean and Mandarin listeners estimated the ages of native Arabic, Korean and Mandarin speakers speaking English.

Our hypotheses are as follows:

- 1a) Listeners will be more accurate and faster at estimating ages of speakers of the same native language.
- 1b) Listeners will be more accurate and faster at estimating ages of speakers with greater linguistic familiarity.
- 2) Speakers of Korean and Mandarin will be perceived as younger than speakers of Arabic. In addition, we will also investigate which acoustic cues contribute to age perception. Specifically, we will test F0 and speech rate.

Hypothesis 1 (a & b) relates to ‘linguistic familiarity’ (see Method). Based on previous findings that experience and familiarity result in more accurate age estimations and faster reaction times (Bürkle and Gnevsheva, 2017; Rodrigues and Nagao, 2010), we hypothesise that listeners with greater linguistic familiarity will estimate age more accurately and faster. Specifically, we expect that when listening to speakers of the same native language (Arabic listeners listening to Arabic speakers, Koreans listening to Koreans, Mandarin listeners listening to Mandarin speakers, i.e., the speaker and listener languages *match*), participants will be faster and more accurate in their age estimation compared to when listening to speakers of the other languages (hypothesis 1a). Korean–Mandarin and Mandarin–Korean are considered to be *close* linguistic familiarity groups. Korean–Arabic, Mandarin–Arabic, Arabic–Korean and Arabic–Mandarin are considered to be *far* linguistic familiarity groups. We expect that accuracy will be highest and reactions fastest for the same-language speaker-listener pairs, in between for Mandarin–Korean pairs and lowest when either the speaker or the listener is Arabic. Listeners with greater linguistic familiarity (the ‘close’ condition) will estimate more accurately and faster than the ‘far’ linguistic familiarity groups, but less so than the ‘match’ group (hypothesis 1b).

Few studies have investigated whether there are qualitative differences in age estimation between speakers of different native languages. One previous study suggests that native (L1) English listeners perceive L1 Japanese speakers of English as younger than native English speakers (Bürkle and Gnevsheva, 2017). However, since this study had a native and a non-native speaker group, it is unclear whether the effect was a language or nativeness effect. We know of no previous study that has explored age perception of different groups of second language speakers. Therefore, in the present study, we investigate whether there are differences in age perception between Arabic, Korean and Mandarin speakers speaking English. Younger voices have higher pitch than older voices (Hartman and Danhauer, 1976; Schötz, 2007; Shipp et al., 1992) and previous research suggests that the mean F0 of a language may affect perception of the age of its speakers (Bürkle and Gnevsheva, 2017). Korean and Chinese have a higher mean F0 than Arabic (Jung et al., 2016; Natour and Wingate, 2009; Zhang et al., 1999). We therefore predict that - if the mean F0 is carried over from the native language to the second language - native Korean and Mandarin speakers would be perceived as younger than native Arabic speakers (hypothesis 2).

Because Mandarin is a tonal language, tonal contours affect F0 trajectories within Mandarin syllables. This differs from Korean and Arabic, which are both non-tonal languages. However, tones are not expected to play a role when using English, except that - as in all languages - there may be carry-over effects of F0 on the pronunciation in English. If this effect is there, we hope to capture at least part of it with the F0 measure.

2. Method

2.1. Linguistic familiarity

To test the effect of linguistic familiarity on age estimation accuracy (hypotheses 1a & 1b), we created a three-level factor of linguistic familiarity (match: speaker L1 is the same as listener L1; close: Korean listeners listening to Mandarin speakers and vice versa; far: Korean or Mandarin listeners listening to Arabic speakers and Arabic listeners listening to Korean or Mandarin speakers; see Table 1).

We note that the term ‘linguistic familiarity’ is used in a broad sense in this study. We expected that Chinese and Korean language speakers would be more familiar with each other’s languages than with Arabic and vice versa. There may be several contributing factors, including geographical proximity of the countries and historical and cultural contact. Although Korean and Chinese are not related languages, approximately 60% of Korean vocabulary was borrowed from Ancient Chinese (Sohn, 2001). China shares a border with the Korean peninsula, and has a long history of social, cultural and economic exchanges with Korea. Diplomatic relations with South Korea were established in 1992 and in

Table 1
Coding of linguistic familiarity between language groups.

Speaker L1	Listener L1	Linguistic familiarity
Arabic	Arabic	Match
Korean	Korean	Match
Mandarin	Mandarin	Match
Korean	Mandarin	Close
Mandarin	Korean	Close
Korean	Arabic	Far
Mandarin	Arabic	Far
Arabic	Korean	Far
Arabic	Mandarin	Far

1999 Korean was already the fourth most popular foreign language in China (Sohn, 2001). Over the last few decades, there has been an increasing level of cultural exchange between Korea and China in terms of television dramas, movies, music and other cultural products (Jang, 2003, 2012; Rowan, 2009). Results of such a close geographical proximity and a relatively high level of cultural exchange between China and Korea may have led to greater linguistic familiarity between the people, compared to Arabic speakers with whom the cultural exchange with China and Korea is still relatively small.

In addition, New Zealand (NZ) census data¹ reveals that there were approximately 5 times as many Chinese speakers (approximately 52,000) and 2.5 times as many Korean speakers (approximately 26,000) as Arabic speakers (approximately 10,700) in NZ in 2013. Therefore, the chance of hearing Arabic speakers speaking English is lower than hearing Mandarin or Korean speakers speaking English.

Our linguistic familiarity assumption was supported by participants’ questionnaire data (see Procedure below). All the listeners were most familiar with their own native language except for two listeners (the match condition). In addition, 70% (14 out of 20) Mandarin listeners were more familiar with Korean language than Arabic language, and 60% (9 out of 15) Korean listeners were more familiar with Mandarin than with Arabic (the close condition). 60% (9 out of 15) Arabic listeners were equally unfamiliar with both Mandarin and Korean languages (the far condition). On a scale from 1 (‘not familiar at all’) to 5 (‘very familiar’), Arabic listeners’ mean linguistic familiarity was 4.8 for Arabic, 1.6 for Korean and 1.7 for Mandarin. Korean listeners’ mean linguistic familiarity was 1.1 for Arabic, 5.0 for Korean and 2.3 for Mandarin. Mandarin listeners’ mean linguistic familiarity was 1.4 for Arabic, 2.2 for Korean and 5.0 for Mandarin. This provides support for our characterisation of Mandarin and Korean speakers being more familiar with each other’s languages than with the Arabic language and vice versa.

2.2. Listeners

Fifty-two participants were recruited as listeners for this experiment. Two participants’ data were excluded due to technical problems during data collection. The data from 50 participants were analysed. They were 15 Arabic² (7 males), 15 Korean (9 males), and 20 Mandarin (5 males) first language speakers. Their ages ranged from 18 to 32 years with the mean age of 24. All participants reported normal hearing. They were recruited in Christchurch and in particular around the University of Canterbury. Participants were given a gift voucher for their participation.

2.3. Stimuli

Ninety-six audio clips were used as stimuli, which were between 22–56 seconds long depending on the speech rate. The stimuli were record-

¹ Statistics New Zealand. Retrieved 8 March 2018 from <http://archive.stats.govt.nz/~media/Statistics/Census/2013%20Census/data-tables/totals-by-topic/totals-by-topic-tables.xls>.

² The Arabic participants in this study were from Palestine, Egypt, Saudi Arabia and Jordan.

Table 2

Mean and standard deviation of accuracy calculated as the absolute difference between real age and estimated age for each language group.

Listeners	Arabic		Korean		Mandarin	
	Mean	SD	Mean	SD	Mean	SD
Speakers						
Arabic	8.30	5.87	9.66	7.58	9.19	6.96
Korean	9.84	7.62	9.35	7.66	9.56	7.19
Mandarin	11.67	9.97	11.79	10.33	11.42	10.25

ings of the same paragraph from 48 speakers, divided into two parts. Sixteen Arabic,³ 16 Korean and 16 Mandarin speakers were selected (age range 21–63 years; mean age 40 years). Forty-two of these clips were extracted from the Speech Accent Archive (Weinberger, 2015), where speakers read the ‘Please Call Stella’ passage in English (see Appendix A). These 42 audio clips were downloaded as .mp3 files and converted to .wav format using Praat (Boersma and Weenink, 2017). All 42 speakers were formally taught English in the classroom, except one who reported having learned English without formal training. All 42 speakers from the Archive were living in the US at the time of the recording.

In order to have a relatively even spread of ages for each group, we aimed to select one female and one male speaker at five-year age intervals between 22 and 62 years of age for each language (namely, ages 22, 27, 32, 37, 42, 47, 52, 57, 62). If an appropriate speaker of that age was not available, we allowed for a deviation of one year (e.g. age 21 or 23).⁴ With these criteria, we were not able to obtain all our speakers from the Speech Accent Archive. Therefore, we supplemented the materials with recordings of six additional speakers (2 Korean, 4 Mandarin).

The six additional speakers were recorded reading the same ‘Please Call Stella’ passage in a soundproof booth at University of Canterbury. The recordings were made on a Tascam HD P2 voice recorder, with settings of 44,100 kHz and 16-bit resolution. The microphone was a Beyer microphone, recorded on a single channel. To our knowledge, none of the speakers had any vocal abnormalities. Apart from one speaker who started to learn English at 24 years of age through self-study, all learned English at school before the age of 16. Speakers signed an informed consent form. None of the speakers participated in the perception experiment as listeners.

2.4. Procedure

Experiment presentation and data collection were conducted using PsychoPy (Peirce, 2009; version 1.85.2). Participants sat in front of a computer screen in a quiet lab room at the New Zealand Institute of Language, Brain and Behaviour. They were informed that they would hear speakers speaking English and were requested to estimate each speaker’s age at the end of the clip by typing a number (exact age in years, not an age range) in the response box. They were not given any information about the language or ethnicity of the speakers or an age range to select from. The clips from all speakers of all languages were randomised for each participant and each audio clip was played only once. Participants’ age estimates and their reaction times were recorded. Reaction time was calculated from the end of the clip to the listener’s decision response, namely, pressing the enter key after typing the estimated age. The experiment consisted of three blocks with breaks in between and

³ The Arabic speakers in the stimuli were from Saudi Arabia, Iraq, Syria, Jordan, Israel and Oman.

⁴ Among the 48 speakers in the stimuli, five did not meet the criterion of a one year deviation. They are a 40 year old male Korean speaker, a 49 year old female Korean speaker, a 50 year old male Mandarin speaker, a 55 year old male Arabic speaker, and a 50 year old male Korean speaker.

lasted approximately 40 min in total. None of the participants reported recognising the speakers.

Following the experiment, participants answered a brief questionnaire on their age, sex, country of birth, ethnicity, language background, familiarity with other languages (a 5-point Likert-style scale), education and experience with English (age of onset, time spent in an English speaking country).

3. Results

3.1. Accuracy

All analyses were conducted after removing errors and null responses (0.71% in total). The age estimation, linguistic familiarity and acoustic analyses were conducted on the remaining 4766 data points. Errors in the data include three types: an age range (e.g., 20–25) instead of a specific age to the year, typographical errors where a non-numeric character was entered (i.e., 27') and age estimation under 5 years old.

Accuracy is calculated as the absolute difference between a speaker’s real age and their estimated age. Therefore, the larger the absolute value, the less accurate the estimate. The mean (and standard deviation, SD) accuracy for all listeners and speakers was 10.08 (8.37). In Nagao (2006), where both English and Japanese participants were used as speakers and listeners, the mean (and SD) of accuracy for all listeners listening to all speakers was 3.77 (18.11) (p. 116). In the present study, the mean (and SD) accuracy for Arabic, Korean and Mandarin listeners was respectively 9.94 (8.12), 10.27 (8.68) and 10.06 (8.32).

Given our hypothesis that linguistic familiarity plays a role in age estimation accuracy, we expect listeners to vary in age estimation accuracy across different speaker-listener groups. The mean and SD of accuracy for three language groups were demonstrated for different speaker - listener L1 combinations, as shown in Table 2.

Numerically, the mean accuracy suggests that, for Arabic speakers, Arabic listeners were the most accurate listener group, followed by Mandarin and Korean listeners. Korean listeners exhibited the highest accuracy for Korean speakers, followed by Mandarin and then Arabic listeners. Mandarin listeners gave the most accurate estimates for Mandarin speakers, followed by Arabic and then Korean listeners. That is, for each speaker group, listeners of the same native language produced the numerically most accurate age estimates. Interestingly, since the overall accuracy for Mandarin speakers was so low, Mandarin listeners were less accurate at estimating Mandarin speakers than they were at estimating Korean and Arabic speakers.

3.2. Linguistic familiarity

To investigate these differences statistically, we fit a linear mixed effects model (Baayen et al., 2008) in R (R Core Team, 2016; version 3.3.2) using the lme4 package (Bates et al., 2015). The dependent variable was accuracy. We tested predictors of trial, linguistic familiarity, speaker age, gender and L1, listener age and gender. Listener L1 was not included because linguistic familiarity is already a combination of speaker L1 and listener L1. Speaker and listener age were log transformed. Trial and speaker age were scaled. Random intercepts of speaker, listener and clip were included. We used a combination of forward and backward model fitting. Each fixed effect and interaction was tested by comparing to the simpler model via an analysis of variance (ANOVA). Random slopes for the significant fixed effects were also individually added and tested. The final model retained only effects that significantly improved model fit.

The final model (see Appendix B for all model formulas) included fixed effects of trial, linguistic familiarity, speaker age, speaker L1 and a two-way interaction between speaker age and linguistic familiarity, random intercepts of speaker, listener and clip, and random slopes of speaker age and speaker L1 for listener. Speaker gender, listener age and listener gender did not improve the model, thus were removed.

Table 3
Model summary of effects of linguistic familiarity on age estimation accuracy.

Fixed effects	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	2.216e-01	2.399e-02	5.700e+01	9.236	< 0.0001
Trial	-7.329e-03	2.335e-03	4.621e+03	-3.139	0.002
Speaker age	8.292e-02	1.513e-02	8.000e+01	5.479	< 0.0001
Familiarity_close	8.303e-03	9.015e-03	5.500e+01	0.921	0.361
Familiarity_far	2.514e-02	8.303e-03	5.800e+01	3.028	0.004
Speaker L1_korean	2.181e-02	3.250e-02	5.300e+01	0.671	0.505
Speaker L1_mandarin	7.403e-02	3.282e-02	5.400e+01	2.255	0.028
Speaker age:familiarity_close	-1.011e-02	6.627e-03	4.580e+03	-1.526	0.127
Speaker age:familiarity_far	-1.413e-02	5.620e-03	4.596e+03	-2.514	0.012

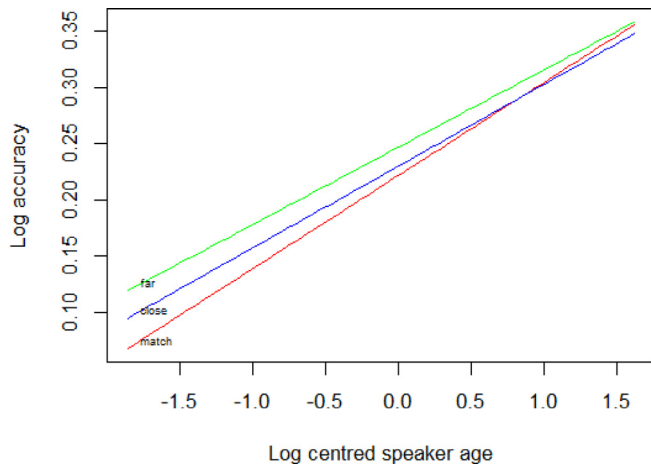


Fig. 1. Log accuracy over speaker age (logged and centred) in the three linguistic familiarity conditions (match, close and far).

The model summary is shown in Table 3. The baseline was Arabic speakers in the match familiarity condition. The model revealed significant effects of trial, speaker age, linguistic familiarity, speaker L1 and an interaction between speaker age and linguistic familiarity. Because the dependent variable, accuracy, was the absolute value of real age minus estimated age, greater coefficients indicate lower accuracy. The effect of trial shows that listeners became more accurate at estimating ages over the course of the experiment. They were less accurate as speaker age increased. There was also a significant effect of linguistic familiarity. Listeners estimated speakers' ages less accurately in the far condition than in the match condition. This finding supports our hypothesis 1b that the listeners would estimate ages more accurately with greater linguistic familiarity. Interestingly, this effect seemed to diminish with speaker age, as indicated by the interaction between speaker age and linguistic familiarity. As illustrated in Fig. 1, the difference in accuracy between far and match conditions was only present for the younger speakers.

However, our hypothesis that listeners would be more accurate with speakers of the same L1 than speakers in the close condition was not supported by the data, as the difference between match and close conditions was not significant ($p = 0.361$). When the model was relevelled, the difference between the close and far conditions was also not significant ($p = 0.107$). The accuracy score was numerically in between the close and far conditions. It is possible that the lack of a significant effect here may be because of a lack of statistical power.

3.3. Reaction time

Hypotheses 1a and 1b involving reaction time (listeners will respond faster to the more familiar languages than the less familiar languages) were tested in a mixed effects model in R with log transformed reaction time as the dependent variable. In addition to removing errors (see Section 3.1), a further two data points were removed, for which lis-

teners' reaction times were over 60 s. Analysis was conducted on the remaining 4764 data points. The fixed effects were continuous predictors of trial, speaker age, listener age and accuracy (the absolute difference between a participant's real age and the estimated age), two-level factors of speaker gender and listener gender and three-level factors of speaker L1 and linguistic familiarity. In the model comparison process, both speaker L1 and linguistic familiarity were found to be non-significant, so listener L1 was tested and revealed to be a significant predictor, thus listener L1 was included in the final model. Speaker age, listener age and accuracy were log transformed. Trial and speaker age were scaled. Speaker, listener and clip were included as random intercepts. Each fixed effect and interaction was tested and compared using an ANOVA. Random slopes were also added for the significant fixed effects. Only the effects that significantly improved model fit remained in the model.

The final model included main effects of trial, speaker gender, listener L1, accuracy and a two-way interaction between speaker gender and listener L1. Random intercepts of speaker, listener and clip were included, as well as random slopes of trial and speaker gender for listener, and a random slope of listener L1 for speaker. No other main effects or interactions were significant, so were removed from the final model.

On the intercept are female speakers with Arabic listeners. The model summary (Table 4) shows that trial, speaker gender, accuracy and the interaction between speaker gender and listener L1 were significant predictors of reaction time. The difference between Arabic and Korean L1 listeners approached significance. The trial effect suggests that listeners became faster over the course of the experiment. The speaker gender effect reveals that Arabic listeners were faster at estimating the age of male speakers compared to female speakers. For the other listener groups, there was no effect of gender. There was a significant effect of accuracy (the absolute difference between real age and estimated age). As reaction time increased, the absolute difference also increased. This suggests an effect of difficulty: as difficulty increased, listeners became both slower and less accurate. The listener L1 effect indicates that Korean listeners tended to be marginally faster at estimating speakers' ages than Arabic listeners. We did not have any prior hypotheses about speed of response for either gender or listener L1 and the present results are rather marginal, so we hesitate to draw any strong conclusions about these effects based on the present results. The reaction time results also do not suggest that linguistic familiarity had an effect on response time in cross-linguistic age estimation. Therefore, the reaction time component of hypothesis 1 (a & b) is not supported.

3.4. Estimated age

In order to test the second hypothesis, that Korean and Mandarin speakers would be perceived as younger than Arabic speakers, a mixed effects model was fit to the data in R with estimated age as the dependent variable. Predictor variables were continuous predictors of trial, speaker age and listener age, and factors of speaker gender, speaker L1, listener gender and listener L1. Estimated age, speaker age and listener age were all logarithmically transformed. Trial was scaled. Random intercepts for speaker, listener and clip were included. Each fixed effect

Table 4
Model summary of effects of speaker and listener characteristics on reaction time.

Fixed effects	Estimate	Std. Error	Df	t value	Pr(> t)
(Intercept)	1.76232	0.07987	51.1	22.065	< 0.0001
Trial	-0.09561	0.01058	49.8	-9.034	< 0.0001
Speaker gender_male	-0.03657	0.01787	396.9	-2.047	0.041
Listener L1_korean	-0.21608	0.11008	50.3	-1.963	0.055
Listener L1_mandarin	-0.17094	0.10291	50.1	-1.661	0.103
Accuracy	0.06685	0.02488	674.8	2.687	0.007
Speaker gender_male: listener L1_korean	0.04591	0.02602	136.8	1.764	0.080
Speaker gender_male: listener L1_mandarin	0.05109	0.02385	212.6	2.142	0.033

Table 5
Model summary of effects of speaker and listener characteristics on age estimates.

Fixed effects	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	2.155830	0.185261	53.58	11.637	< 0.0001
trial	0.018253	0.004593	49.92	3.974	< 0.001
speaker age	0.387243	0.050348	54.01	7.691	< 0.0001
speaker L1_korean	-0.048865	0.037555	52.49	-1.301	0.199
speaker L1_mandarin	-0.122141	0.038205	55.71	-3.197	0.002

Table 6
The acoustic measurements of speakers per gender per language.

Speaker L1	Gender	Mean F0 (Hz)	Speech Rate (syll/s)
Arabic	Female	211.21	2.08
	Male	116.80	2.33
Korean	Female	188.94	2.59
	Male	129.22	2.73
Mandarin	Female	215.78	2.33
	Male	132.16	2.32

and interaction was tested. A random slope was individually added and tested for each of the significant fixed effects. An ANOVA was used for model comparisons. Only predictors that significantly improved model fit were retained in the final model.

The best fit model included fixed effects of trial, speaker age and speaker L1, with no interactions (see Table 5). Random intercepts of speaker, listener and clip were included, as well as random slopes of trial, speaker age and speaker L1 for listener. Speaker gender, listener age, listener gender, listener L1 and the two-way interactions between speaker and listener L1 did not improve model fit, so were removed from the model.

In the model summary, the Arabic L1 level of speaker is on the intercept. Trial, speaker age and speaker L1 were significant predictors of estimated age. The trial effect shows that age estimates increased over the course of the experiment. The effect of speaker age indicates that older speakers were estimated to be older than the younger speakers, but that age increases were underestimated. More importantly for our second hypothesis, age estimates for Mandarin L1 speakers were significantly lower than those for Arabic speakers (see Fig. 2). This result provides partial support for our second hypothesis, namely that Mandarin speakers are perceived as younger than Arabic speakers. The difference between Korean L1 speakers and Arabic L1 speakers was not significant.

3.5. Acoustic analysis

The above results reveal several sociolinguistic factors that affect age estimation. Here we investigate the question of which acoustic cues listeners may have been attending to in their estimation decisions. We measured speakers' mean F0 in Hz and their average speech rates in syllable-per-second (generated by dividing speech duration by syllable count). The measurements are presented in Table 6. Numerically, it ap-

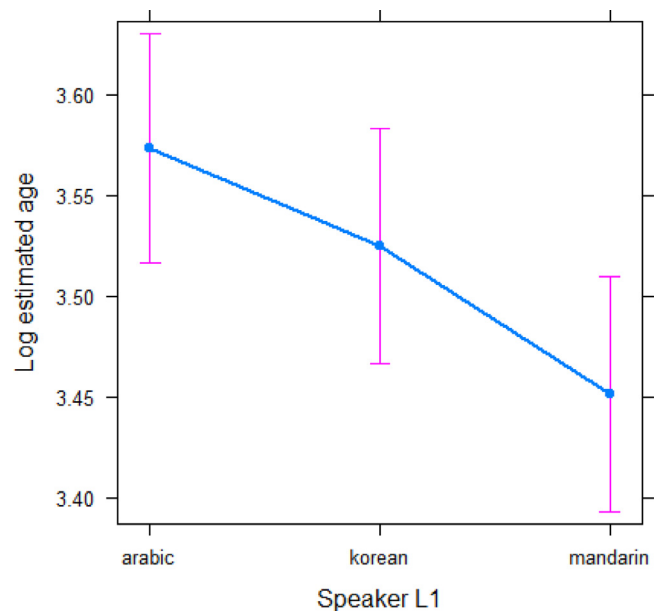


Fig. 2. Log estimated age of Arabic, Korean and Mandarin L1 speakers.

pears that both Korean and Mandarin male speakers had a higher mean F0 than Arabic male speakers, but for the female speakers, Koreans had the lowest mean F0. Mandarin speakers had the highest mean F0 in both males and females. In terms of speech rate, Arabic speakers in both males and females were numerically the slowest group. Korean speakers were faster than Mandarin speakers regardless of the gender.

To test whether these acoustic cues affected age estimation, we fit a mixed effects model to the acoustic data. The acoustic cue model fitting procedure was similar to the analyses of sociolinguistic factors in age estimation above. The dependent variable was estimated age. We tested predictors of trial, speaker age, gender and L1, acoustic factors of mean F0 and speech rate, as well as linguistic familiarity. Random intercepts included speaker, listener and clip. Estimated age, speaker age and mean F0 were all log transformed. Trial was scaled. Each main effect and interaction was tested via an ANOVA comparison. A random slope was individually added for the significant fixed effects.

The final model included fixed effects of trial, speaker age, speaker L1, mean F0, speech rate and linguistic familiarity without any interactions, random slopes of trial, speaker age, speaker L1, mean F0 and speech rate for listener, and random intercepts of speaker, listener and clip. Speaker gender and all the interactions did not improve the model fit, so were removed.

The model summary is presented in Table 7. Apart from the significant effects of trial, speaker age, speaker L1, linguistic familiarity on age estimates discussed previously, acoustic factors of mean F0 and speech rate also had significant effects on age estimates. The effect of mean F0 indicated that when the mean F0 of speakers was lower, they were estimated to be older by the listeners (see Fig. 3). The effect of speech

Table 7
Model summary of effects of acoustic cues on age estimates.

Fixed effects	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	3.299615	0.346479	58.57	9.523	< 0.0001
trial	0.018062	0.004495	49.93	4.018	< 0.001
speaker age	0.323011	0.051320	53.21	6.294	< 0.0001
speaker L1_korean	-0.020159	0.037295	54.98	-0.541	0.591
speaker L1_mandarin	-0.099027	0.037214	57.00	-2.661	0.010
familiarity_close	0.018599	0.012407	50.94	1.499	0.140
familiarity_far	0.023508	0.010992	58.02	2.139	0.037
mean f0	-0.156693	0.051855	67.88	-3.022	0.004
speech rate	-0.060409	0.018336	8.32	-3.295	0.010

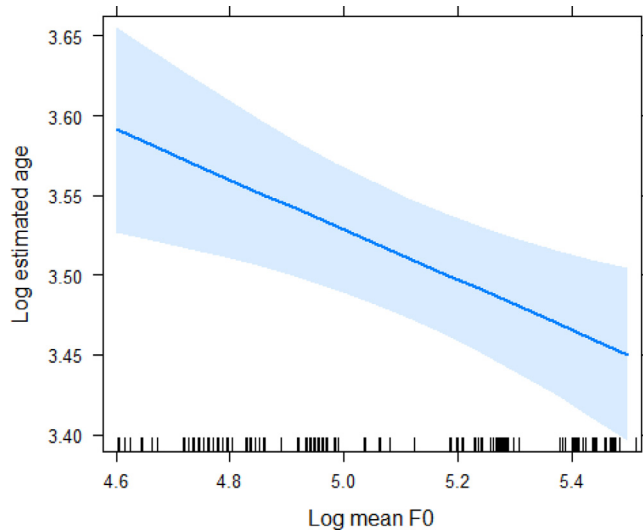


Fig. 3. The effect of log mean F0 on log estimated age.

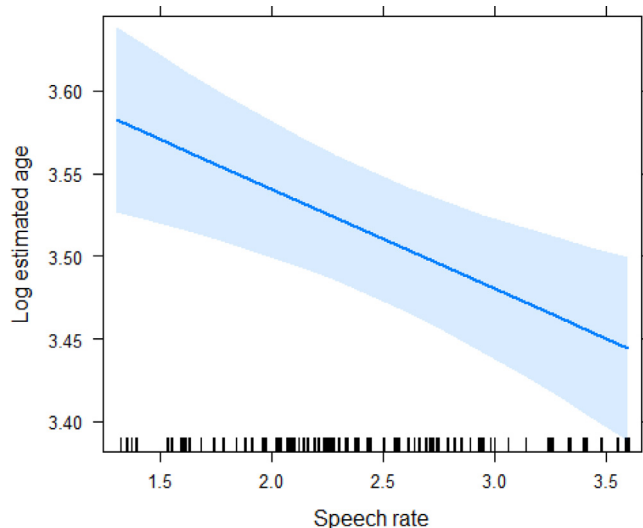


Fig. 4. The effect of speech rate (syllable/second) on log estimated age.

rate showed that the faster the speaker read the stimulus passage, the younger they were estimated to be by the listener (see Fig. 4). The effect of linguistic familiarity or speaker L1 did not interact with either mean F0 or speech rate in this study. This suggests that mean F0 had the same effect across the language groups.

It is worth noting that when F0 was added to the baseline model, it did not significantly predict the age estimate. Only when other predictors were present in the model, most notably speaker age, did mean F0 become a significant predictor. This demonstrates that listeners are not

relying solely on F0 for their age estimates. Listeners must instead rely on at least one, probably several, other acoustic cues to make their estimates. In combination, F0 had a significant effect on the age estimate.

4. Discussion and conclusion

This study investigated English as a second language (L2) listeners' estimation of English L2 speakers' age. Native Arabic, Korean and Mandarin participants listened to recordings of native Arabic, Korean and Mandarin speakers speaking English and estimated their age to the year. We expected a) age estimation accuracy to decrease and reaction times to increase with decreasing linguistic familiarity and b) Mandarin and Korean speakers to be perceived as younger than Arabic speakers.

The linguistic familiarity model showed that when the listener and speaker had the same L1, age estimation was more accurate than when the speaker and listener had low linguistic familiarity. These findings demonstrate that linguistic familiarity can affect listeners' age estimation, with higher accuracy when listening to speakers of one's own L1, even when both speaking and listening are in the L2. This finding is in line with and extends the findings of [Rodrigues and Nagao \(2010\)](#), who reported that age estimation was better in native English listeners who were more familiar with foreign accented English.

The present study did not find evidence for our hypothesis that listeners estimate age more accurately in the 'match' condition (where listeners and speakers have the same L1) than the 'close' condition (Korean listeners estimate Mandarin speakers and vice versa). Nor was our hypothesis that listeners estimate age more accurately in the 'close' condition (Korean listeners estimate Mandarin speakers and vice versa) than the 'far' condition (Korean and Mandarin listeners estimate Arabic speakers, and Arabic listeners estimate Korean and Mandarin speakers) supported by this study. As these are null effects, no strong conclusions can be drawn. Our failure to find differences between these groups may be due to a lack of sufficient power in the present experiment. However, it is worth noting that previous studies have also failed to find differences in age estimation accuracy between closely related European languages. [Braun and Cerrato \(1999\)](#) failed to find an effect of native language in an Italian-German age estimation study, even when the participants spoke their native languages. In addition, this study cannot provide evidence for a relationship between reaction time and linguistic familiarity. This may be because in our experiment, participants had to wait until the audio clip had finished before responding. This may have levelled any differences between familiarity conditions in response time.

Our second hypothesis of age estimation was partially supported. The estimated age model indicated that Mandarin speakers were perceived to be younger than Arabic speakers. However, the hypothesis that Korean speakers would be perceived to be younger than Arabic speakers was not supported in this study.

Our acoustic model showed that fundamental frequency and speech rate were significant predictors of age estimation of non-native speakers by non-native listeners, if speaker age was also included in the model. The significant effect of mean F0 is in line with [Bürkle and Gnevsheva's \(2017\)](#) finding that speakers with a higher F0 are perceived to be younger. In addition to F0, the acoustic cue of speech rate was also significant. The effect of speech rate suggests that speakers with a slower rate of speech are estimated as older than quicker speaking ones, irrespective of listener background. Previous studies have shown that speech rate slows with age ([Harnsberger et al., 2010](#)). However, the speech rate in L2 speech also tends to be lower at lower proficiencies and was relatively low in our data. Therefore, it might be expected that variability due to proficiency might override effects of speech rate on age estimation. However, our data suggests that speech rate affects age perception even in L2 speech. More linguistic features such as tremor, language proficiency etc. can be considered in the future investigations of age perception cues in foreign accented speech.

The effect of trial showed that listeners became both quicker (reaction time model, [Table 4](#)) and more accurate (accuracy model, [Table 3](#))

over the course of the experiment. This suggests a learning effect, such that with more experience of different voices, listeners got better at discriminating between different ages, even in the absence of explicit feedback about the accuracy of their responses.

Listeners were less accurate when listening to older speakers. In the present study, we restricted the age of listeners to the 18–32 range to reduce variability. Therefore, we cannot tease apart whether this was due to a main effect of speaker age or difference between listener age and speaker age. However, the results are consistent with previous findings of an ‘own-age bias’, where listeners estimate ages more accurately when speakers are similar to their own age (Moyses, 2014). For these young listeners, estimating older speakers’ ages was more difficult than estimating the ages of peers. This effect may stem from voice cue familiarity. Particularly in an educational setting, the majority of voices that young people hear around them are likely to be younger. Students have a lot of experience with similar-age students and groups of small age increments. In addition, there may be more social significance attached to small age differences for younger speakers, compared to older speakers. For example, in a university setting, a few years can mean the difference between a first year student, a final year student or a post-graduate student.

The effect of linguistic familiarity in the present study suggests that during age estimation, listeners are not only sensitive to cues affected by the biological changes that occur with increasing age, but are also affected by language-specific social-cultural traits of speakers (also cf. Bürkle and Gnevshva, 2017; Schötz, 2007). The effect of linguistic familiarity may have practical implications for age estimation in real life situations, such as forensics. Age estimates are often used in forensic settings. If witnesses only have information about a perpetrator’s voice without seeing the perpetrator, age estimates can only be obtained from speech. Witnesses’ testimony might then be influenced by the perpetrator’s and witness’s native languages and may be less accurate for unfamiliar language backgrounds, even if the language spoken is English.

Acknowledgements

This work was supported by a ‘Language Variation and Change’ grant to all authors from the New Zealand Institute of Language, Brain and Behaviour. We are grateful to Dr Kevin Watson and Professor Jen Hay for contributions to experiment design and feedback on an earlier version of this paper. We also thank Moonsun Choi, Yoonmi Oh, Mohammed Dagamseh and Muneir Gwasmeh, who provided their help with recruiting Korean and Arabic participants. We thank all the individuals who participated in this study. Finally, we thank two anonymous reviewers for their valuable feedback. Any remaining errors in this article are our responsibility.

Appendix A. ‘Please Call Stella’ passage

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

Appendix B. Formulas for models

B.1 Linguistic familiarity: accuracy \sim trial + speaker_age*linguistic_familiarity + speaker_L1 + (1 + speaker_age + speaker_L1|listener) + (1|speaker) + (1|clip)

B.2 Reaction time: reaction time \sim trial + speaker_gender*listener_L1 + accuracy + (1 + trial + speaker_gender|listener) + (1 + listener_L1|speaker) + (1|clip)

B.3 Estimated age: estimated age \sim trial + speaker_age + speaker_L1 + (1 + trial + speaker_age + speaker_L1|listener) + (1|speaker) + (1|clip)

B.4 Acoustic cues: estimated age \sim trial + speaker_age + speaker_L1 + linguistic_familiarity + mean_f0 + speech_rate + (1 + trial + speaker_age + speaker_L1 + mean_f0 + speech_rate |listener) + (1|speaker) + (1|clip)

References

- Amilon, K., van de Weijer, J., Schötz, S., 2007. The impact of visual and auditory cues in age estimation. In: Müller, C. (Ed.), *Speaker Classification II: Selected Projects*. Springer, Berlin, pp. 10–21. doi:10.1007/978-3-540-74122-0_2.
- Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modelling with crossed random effects for subjects and items. *J. Memory Lang.* 59, 390–412.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1), 1–48.
- Boersma, P., Weenink, D., 2017. Praat: doing phonetics by computer. Version 6.0.36 Retrieved 31 July 2017 from.
- Braun, A., 1996. Age estimation by different listener groups. *Forensic Linguist.* 3 (1), 65–73.
- Braun, A., Cerrato, L., 1999. Estimating Speaker Age Across Languages. *ICPhS99*, pp. 1369–1372. Retrieved 19 December 2017 from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_1369.pdf.
- Bürkle, D., Gnevshva, K., 2017. Age estimation in foreign-accented speech. In: *Proceedings of the International Conference on Language Variation in Europe (ICLaVE)*. Malaga, Spain.
- Cerrato, L., Falcone, M., Paoloni, A., 2000. Subjective age estimation of telephonic voices. *Speech Commun.* 31, 107–112. doi:10.1016/S0167-6393(99)00071-0.
- Harnsberger, J.D., Brown, W.S., Shrivastav, R., Rothman, H., 2010. Noise and tremor in the perception of vocal aging in males. *J. Voice* 24 (5), 523–530.
- Harnsberger, J.D., Shrivastav, R., Brown, W.S., Rothman, H., Hollien, H., 2008. Speaking rate and fundamental frequency as speech cues to perceived age. *J. Voice* 22 (1), 58–69.
- Hartman, D.E., 1979. The perceptual identity and characteristics of aging in normal male adult speakers. *J. Commun. Disord.* 12, 53–61. doi:10.1016/0021-9924(79)90021-2.
- Hartman, D.E., Danhauer, J.L., 1976. Perceptual features of speech for males in four perceived age decades. *J. Acoust. Soc. Am.* 59 (3), 713–715. doi:10.1121/1.380894.
- Hollingworth, H.L., 1910. The central tendency of judgment. *J. Philos. Psychol. Sci. Methods* 7 (17), 461–469. doi:10.2307/2012819.
- Hughes, S.M., Rhodes, B.C., 2010. Making age assessments based on voice: the impact of the reproductive viability of the speaker. *J. Soc. Evolut. Cultural Psychol.* 4 (4), 290–304. doi:10.1037/h0099282.
- Huntley, R., Hollien, H., Shipp, T., 1987. Influences of listener characteristics on perceived age estimates. *J. Voice* 1 (1), 49–52.
- Jang, S.H., 2003. Contemporary Chinese narratives on Korean culture. *Korea J.* 43 (1), 129–153.
- Jang, S.H., 2012. The Korean wave and its implications for the Korea–China relationship. *J. Int. Area Stud.* 19 (2), 97–113.
- Jung, Y.-J., Lim, S.B., Lee, G., Rhee, S.-C., 2016. Comparing speaking fundamental frequency in bilingual and L2-accented speech. *J. Acoust. Soc. Am.* 140 (4), 3270–3270. doi:10.1121/1.4970381.
- Krauss, R.M., Freyberg, R., Morsella, E., 2002. Inferring speakers’ physical attributes from their voices. *J. Exp. Soc. Psychol.* 38, 618–625. doi:10.1016/S0022-1031(02)00510-3.
- Linville, S.E., 1987. Acoustic-perceptual studies of aging voice in women. *J. Voice.* 1 (1), 44–48.
- Linville, S.E., Korabic, E.W., 1986. Elderly listeners’ estimates of vocal age in adult females. *J. Acoust. Soc. Am.* 80 (2), 692–694. doi:10.1121/1.394013.
- Moyses, E., 2014. Age estimation from faces and voices: A review. *Psychol. Belgica* 54 (3), 255–265. doi:10.5334/pb.aq.
- Moyses, E., Beaufort, A., Brédart, S., 2014. Evidence for an own-age bias in age estimation from voices in older persons. *Eur. J. Ageing.* 11, 241–247. doi:10.1007/s10433-014-0305-0.
- Nagao, K., 2006. Cross-Language Study of Age Perception (Ph.D.). Indiana University Retrieved 26 December 2017 from <http://search.proquest.com/docview/305335298>.
- Nagao, K., Kewley-Port, D., 2005. The effect of language familiarity on age perception. In: *Proceedings of the International Research Conference on Aging and Speech Communication*. Bloomington, USA.
- Natour, Y.S., Wingate, J.M., 2009. Fundamental frequency characteristics of Jordanian Arabic speakers. *J. Voice* 23 (5), 560–566. doi:10.1016/j.jvoice.2008.01.005.
- Peirce, J.W., 2009. Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform.* 2 (10). doi:10.3389/neuro.11.010.2008.
- Ptacek, P.H., Sander, E.K., 1966. Age recognition from voice. *J. Speech Hear. Res.* 9, 273–277. doi:10.1044/jshr.0902.273.
- R Core Team, 2016. R Foundation for Statistical Computing, Vienna, Austria URL <http://www.R-project.org/>.
- Rodrigues, P., Nagao, K., 2010. Effects of listener experience with foreign accent on perception of accentedness and speaker age. *J. Acoust. Soc. Am.* 127 (3), 1956. doi:10.1121/1.3384968.
- Rowan, P., 2009. Korean pop music in China: nationalism, authenticity, and gender. In: *Berry, C., Liscutin, N., Mackintosh, J.D. (Eds.), Cultural Studies and Cultural Industries in Northeast Asia: What a Difference a Region Makes*. Hong Kong University Press, Hong Kong, pp. 151–168.
- Ryan, W.J., Burk, K.W., 1974. Perceptual and acoustic correlates of aging in the speech of males. *J. Commun. Disord.* 7, 181–192. doi:10.1016/0021-9924(74)90030-6.

- Schötz, S., 2005. Effects of stimulus duration and type on perception of female and male speaker age. In: *Proceedings of the Fonetik*. Department of Linguistics, Göteborg University, pp. 87–90.
- Schötz, S., 2007. Acoustic analysis of adult speaker age. In: Müller, C. (Ed.), *Speaker Classification I: Fundamentals, Features, and Methods*. Springer, Berlin, pp. 88–107.
- Shipp, T., Hollien, H., 1969. Perception of the aging male voice. *J. Speech Hear. Res.* 12, 703–710. doi:10.1044/jshr.1204.703.
- Shipp, T., Qi, Y., Huntley, R., Hollien, H., 1992. Acoustic and temporal correlates of perceived age. *J. Voice* 6 (3), 211–216.
- Sohn, H.-M., 2001. *The Korean Language*. Cambridge University Press, Cambridge.
- Torre, P., Barlow, J.A., 2009. Age-related changes in acoustic characteristics of adult speech. *J. Commun. Disord.* 42, 324–333.
- Waller, S.S., Eriksson, M., Sörqvist, P., 2015. Can you hear my age? Influences of speech rate and speech spontaneity on estimation of speaker age. *Front. Psychol.* 6, 1–11. doi:10.3389/fpsyg.2015.00978.
- Weinberger, S., 2015. *Speech Accent Archive*. George Mason University Retrieved 12 July 2017 from <http://accent.gmu.edu>.
- Zhang, J., Huang, M., Li, M., Liu, Q., 1999. Analysis of results of fundamental frequency and voice parameter in healthy young people. *Lin Chuang Er Bi Yan Hou Ke Za Zhi* 13 (9), 403–405 (*J. Clinical Otorhinolaryngology*).