

DOES MEDIA EXPOSURE TO AN ACCENT IMPACT UPON THE ESTIMATION OF THE AGE OF SPEAKERS?

Kirk P.H. Sullivan*, Donn Bayard⁺, Ann Weatherall[†], Cindy Gallois[‡], Frank Schlichting[°] & Jeffery Pittam[‡]
*Umeå University, Sweden, [†]University of Otago, NZ, [‡]Victoria University of Wellington, NZ, [‡]University of Queensland, Australia, [°]University of Tübingen, Germany

ABSTRACT: World English accents are beamed via the media into the homes of those living in English speaking countries. The situation is the same in non-English speaking countries with a subtitling rather than dubbing policy. The degree to which an individual is exposed to the range of world Englishes, thus, varies from country to country. Research into the ability of individuals to identify the English accent of a speaker has demonstrated variation, which is dependent upon the country in which one lives and most likely the degree of exposure to non-local world English accents. This paper explores whether exposure to world English accents impacts upon the forensically important question of an individual's ability to assign the perceived age of speakers based on vocal cues alone. 150 English speakers, 50 from New Zealand, 50 from Australia and 50 from the United States of America responded to a range of questions about two New Zealand, two Australian and two North American voices. Only one of the questions related subjective age estimation. Statistical analysis showed no significant differences due to country or residence and hence media exposure to non-local world Englishes. A second set of 150 listeners, 50 from Germany, 50 from Finland and 50 from Sweden undertook the same task. No significant difference was found due to subtitling as opposed to dubbing policy. This study found no impact of media exposure to accent upon the estimation of the age of speakers based on vocal cues alone.

INTRODUCTION: THE MEDIA

Through contemporary media native English speakers are exposed to world English accents other than their own. This is particularly true of the non-North American English speaker who is exposed to large doses of North American English on television and through the cinema. The speaker of North American English receives, for example, far less exposure to New Zealand or Australian English than the New Zealander or Australian English speaker receives of North American English. Although precise figures for the percentage of programmes screened in each country, which are made and screened in that country, are unavailable, well-motivated approximations can be found in Norris et al. (1999). In their study of television in ten countries, they found that in the USA local programming accounted for approximately 90%, whereas, in Australia local programming accounted for approximately 55% and in New Zealand only 24%. The extent of American programming in New Zealand was demonstrated in the mid-1980s by Bayard (1987). He showed that over half of the programmes screened at that time were American ones. It is, thus, clear that the media defines, for the majority, the degree to which an individual is exposed a non-local world English accent.

The variation in exposure to non-local world English accents is the most likely explanation for the differences in the ability of individuals to identify the accent of a speaker, which is reported in Bayard et al. (2001). With the exception of a single case (the most popular guess of accent for the Australian male voice was New Zealand), the majority of New Zealand and Australian listeners in the Bayard et al. study correctly identified the accent of the speaker. This stands in direct contrast to the US listeners who successfully only identified the North American accents correctly. The most popular guesses for both of the British English speakers were correct, but they did not achieve a majority.

In non-English speaking countries, a comparable situation, in which the media controls an individual's exposure to world English accents, is to be found. At the fundamental level a country's media's policy on subtitling, as opposed to dubbing, dictates whether the majority of a country's citizens have daily exposure to English or not. In countries like France, Germany and Austria there is little, or no, exposure to any foreign languages through the media as foreign language cinema and television are dubbed into the national language. This situation contrasts diametrically with that found in countries like the Netherlands, Sweden and Finland, where foreign language cinema and television are broadcast in the foreign language with national language subtitles. The Norris et al. (1999) found that

in Finland local programming accounted for 55% of television screened and that in Norway the figure was 56%. Of the remaining circa 45% of television screened in these countries, the overwhelming majority of programming is from English speaking countries, the majority of which originates in North America. In order to assess the impact of this degree of exposure to English through the media in the Nordic countries, Bayard and Sullivan (2000) undertook a study comparing New Zealand and Swedish listeners' abilities to perceive an English speaker's country of origin. The Bayard and Sullivan study compared the responses to the stimuli used in Bayard (2000) New Zealand study with Swedish university student responses. It was found that "Swedish and New Zealand listeners are similarly able to perceive the country of origin of non-NZ speakers of English", that "although there was an agreement in the plurality for each speaker, the Swedish responses were generally more widely spread across possible countries of origin" and that "as expected NZ listeners are better at identifying NZ voices". This result confirmed the findings of Sullivan and Karst (1996) study comparing the perception of English accent by Swedish university students and British English High school students.

The impact of media exposure to non-local English accents (all English accents are naturally non-local accents in non-English speaking countries) upon an individual's ability to identify a speaker's country of origin has been shown by Bayard et al (2001) in the first language setting and by Bayard and Sullivan (2000) in the second language setting. It can, thus, be assumed that the listener in both the first and the second language context attends to the accent of the programme and not just the programme's content *per se*. The work presented here investigates whether familiarity with a world English accent received primarily via the media results in *differences* in the accuracy of subjective age estimation based on vocal cues alone.

BACKGROUND: SUBJECTIVE AGE ESTIMATION

Listeners are particularly adept at recognizing people by their voices alone. The reliability of this apparent ability has been widely investigated even under extreme conditions (e.g. Rose & Duncan 1995, Schlichting & Sullivan 1996, 1997). In a recent paper Sullivan & Schlichting (2000) investigated the development of speaker discrimination in a foreign language (Swedish) when the language is being learnt in their first language environment. Their study found a difference between those with no knowledge and those with some knowledge of Swedish, yet concluded "the degree of competence in the second language required ... to produce reliable results is higher than that acquired on a four-year BA degree from a British university" (Sullivan & Schlichting: 108). A crime victim may base important descriptive features of the criminal upon voice features. One such feature could be the criminal's age.

Several studies have investigated the ability of listeners to estimate a given speaker's age. Some of these studies have used the correlations between the calendar age and the perceived age (e.g. Neiman & Applegate, 1990), and others have used percentage correctly identified age groups to demonstrate the degree of listener accuracy in subjective age estimation (e.g. Cerrato, Falcone & Paoloni, 1998). All studies prior to Braun and Cerrato (1999) used same-language speakers and listeners. Braun and Cerrato (1999) used German and Italian speakers and listeners — the listeners had no knowledge of the other language. Absolute error, which is expressed as the absolute difference in years between the perceived age and the calendar age, was used by Braun and Cerrato (1999) to establish listener accuracy. They found that "Italian listeners could not cope at all with the stimuli presented to them in German, whereas the German listeners did just as well on the Italian stimuli as did the Italians" (Braun and Cerrato, 1999: 1371). In terms of the average individual error in years (each listener assigned the age to 20 Italian and 20 German speakers), the Italian listeners assigned age with an average error of 8.5 years to the Italian speakers, whereas the German listeners assigned age with an average error of 10.2 years. For the German speakers age was judged with an average error of 10.5 years by the German listeners and 12.1 years by the Italian listeners. Interestingly, none of these differences was statistically significant and thus hints that, unlike speaker discrimination, language competence does not impact upon subjective age assignment based on vocal cues alone.

THE INVESTIGATION

Previous research has shown (Bayard, 2000; Bayard et al. 2001; Bayard and Sullivan, 2000) that listeners are clearly aware of the voices of non-local English speakers, to which they are exposed through the media. Further a range of studies have shown that listeners are able, within tolerable limits, to estimate the age of a speaker from vocal cues alone (e.g. Cerrato, Falcone & Paoloni, 1998,

Braun and Cerrato, 1999). This investigation does not aim to address how accurate the listeners are in their subject estimation of an individual's age, *per se*, but asks the following three primary questions:

1. Is there a difference in subjective age assignment between native English listeners assigning age to speakers of their own English and native English listeners assigning age to speakers of an English accent which is not their own?
2. Can any difference found be attributed to media exposure to non-local world Englishes?
3. Is there a difference in subjective age estimation in the learner of English as a second language due to a country's policy towards subtitling/dubbing?

These questions are of forensic interest: it is important to know whether exposure to non-local world Englishes will impact the assignment of perceived age. Equally, if a crime victim is a non-native speaker of English, yet comes from a country with a policy of dubbing, will they assign perceived age differently to someone from a country whose media uses subtitles? The degree of importance a crime investigator should attribute to perceived age assignment from individuals with different linguistic backgrounds could prove crucial to the success of a crime investigation.

METHOD

The subjective age estimation data from New Zealand, Australian and North American listeners analyzed and discussed in this paper formed one aspect of the data collected and reported in Bayard et al (2001). The method used to gather the additional data reported in this paper is identical to that reported in Bayard et al (2001) and is reported in full in this paper. The work presented here uses only a subset of the voices presented to the listeners during data collection. The listeners' responses to the two UK speakers' voices are not analyzed for subjective age assignment. Further, the question asking the listener to assign the speaker's age is one of many questions the listeners are asked to respond to as part of the data collection process. In a real-life situation it is unlikely that the listener will be solely attending to a speaker's and no other aspects of the speaker's voice or the context/setting in which the voice is heard. The method of obtaining age estimations as part of a larger set of questions must be seen as increasingly the forensic validity of the task.

Experimental participants: the speakers

Eight speaker's accents from four English speaking countries were used as stimuli. There were two speakers from New Zealand, two from Australia, two from North America and two from the UK. In each country pair there was a male and a female speaker. As the responses to the UK voices were not used in this analysis carried out for this study, the details relating to these two voices are not taken up here. They can, however, be found in Bayard et al (2001) as can more detailed information about the phonological features of all eight voices. All speakers were recorded reading a short text developed at the University of Canterbury, NZ (Weatherall et al. 2000). The average duration of the reading was 30.3 sec, and the standard deviation was 2.643. The New Zealand English and Australian English speakers were judged by Bayard and Weatherall to be characteristic of middle-of-the-road General New Zealand and General Australian English accents. The North American voices were judged by native speakers Bayard and George Ray (Cleveland State University) to be typical of the Inland North American English accent. Table 1 shows the speakers' ages at the time of the recordings were made.

Speaker	Age
New Zealand Male	40
New Zealand Female	46
Australian Male	60
Australian Female	38
North American Male	43
North American Female	43

Table 1. The ages of the six English speakers

Experimental participants: the listeners

To address questions 1 and 2, the participants were: 50 New Zealand Otago and Victoria University students randomly selected from a sample of 257 students (male=10, female=40, mean age=21.9); 50 Australian students from the University of Queensland randomly selected from a sample of 99 students (male=20, female=30, mean age=19.4); and 50 North American students from Cleveland

State University, randomly selected from a sample of 53 students (male=18, female=32, mean age=25.1).

To address question 3, the participants were: 50 Swedish students from Umeå University selected randomly from a sample of 115 students (male=16, female=34, mean age=25.4); 50 Finnish students studying at the University of Helsinki selected randomly from a sample of 97 students (male=3, female=46; unreported gender=1, mean age=22.3); and 50 German students studying at the University of Tübingen (male=25, female=25, mean age=23.2).

Experimental procedure

The experiment consisted of two parts; both parts began with a practise trial. The presentation of the voices was the same in both parts, but the order, *per se*, was random. The listeners, thus, heard and responded to one practise voice, the responses to which were not analysed, and eight randomly presented voices, twice. In the first part the participants rated their impressions of each speaker on 13 six-point semantic differential scales with anchors of "not at all" to "very". They also rated the person's voice on five six-point differential scales. In the second part the listeners were asked to indicate the perceived age, ethnicity, education level, occupation, annual income and social class on a set of fixed options. The questions were translated into Swedish, Finnish and German by native speakers.

The set of fixed options for subjective age estimation were 18–25, 26–35, 36–45, 46–55, 56–65 and 65+ years. The limitation of using age groups was pointed out by Braun and Cerrato (1999) drawing on the research of Cerrato, Falcone and Paoloni (1998) wrote: "Age groups [...] will always bear the problem of where to draw the line between groups, i.e. shifting the line by as little as one year may make a vast difference in the percentage of correct assignments." However, as the aim of this paper is not access how accurate the listeners are, *per se*, but rather to assess whether the various groups differ in their responses, and as the listeners were being asked to make a range of other decisions simultaneously that age group data would be sensitive enough to show up any significant differences in group responses.

RESULTS

A univariate analysis of variance was conducted with perceived age as the dependent variable and the variables local/non-local and country of listener residence as fixed factors. There was no significant impact of local versus non-local English at the .05 level ($F_{1,884} = 3.739$; $p = .053$). The accent of the listeners (New Zealand, Australia and the USA) equally proved to have no significant impact upon perceived age assignment ($F_{2,883} = 2.111$; $p = .122$).

In order to examine the data in more detail an analysis of variance was conducted with perceived age of US voices, perceived age of Australian voices and perceived age of New Zealand voices as three dependent variables and the country of listener residence as a fixed factor. No significant difference in response due to country of residence was found for the New Zealand voices ($F_{2,291} = .6$; $p = .549$) or for the Australian voices ($F_{2,294} = 1.970$; $p = .141$). A significant difference was, however, found for the North American voices ($F_{2,293} = 7.556$; $p = .001$). Tukey and Scheffé posthoc tests were applied to the dependent variable perceived age of the US voices. Both tests found significant mean differences at the $p = .05$ level between US and New Zealand listeners, and the US and the Australian listeners. No significant mean difference was found between the New Zealand and the Australian listeners.

Analysis of variance was also conducted at the individual voice level, that is the perceived age for each voice was a separate dependent variable. No significant difference in response due to country of residence was found for the Australian female voice ($F_{2,126} = .984$; $p = .376$), or the New Zealand male voice ($F_{2,143} = 2.620$; $p = .076$). Significant differences at the $p = .05$ level were found for the remaining four voices: the New Zealand female voice ($F_{2,145} = 3.847$; $p = 0.024$), the Australian male voice ($F_{2,145} = 3.941$; $p = .022$), the US female voice ($F_{2,145} = 3.560$; $p = .031$) and the US male voice ($F_{2,145} = 4.418$; $p = .014$). Tukey and Scheffé posthoc tests were applied to the dependent variable perceived age for each of the voices for which the analysis of variance had shown to be significant. Both tests found significant mean differences at the $p = .05$ level between US and New Zealand listeners, and the US and the Australian listeners for all of the four voices. No significant mean difference between the New Zealand and the Australian listeners was found for any of the four voices. In the case of the New Zealand female and Australian male voices the US listeners perceived the voices to be younger than

the New Zealand and Australian listeners. For the US voices the opposite was the case. This perhaps predictably made the US listeners more accurate for the US voices and less accurate for the New Zealand female and Australian male voices than the New Zealand and Australian listeners.

A parallel set of statistical analyses were conducted on the data collected in non-English speaking countries. A univariate analysis of variance was conducted with perceived age as the dependent variable and the variables dubbed/subtitled and first language of the listeners as fixed factors. No significant impact of dubbing versus subtitled of English language programming in non-English speaking countries was found at the .05 level ($F_{1,888} = 2.424$; $p = .120$). The first language of the listeners (German, Swedish or Finnish) equally proved to have no significant impact upon perceived age assignment ($F_{2,888} = 1.373$; $p = .239$).

In order to examine the data in more detail an analysis of variance was conducted with perceived age of US voices, perceived age of Australian voices and perceived age of New Zealand voices as three dependent variables and the country of listener residence as a fixed factor. No significant difference in response due to country of residence was found for the New Zealand voices ($F_{2,293} = .803$; $p = .803$) or for the Australian voices ($F_{2,293} = 1.027$; $p = .106$). A significant difference was, however, found for the North American voices ($F_{2,293} = 4.306$; $p = .014$). Tukey and Scheffé posthoc tests were applied to the dependent variable perceived age of the US voices. Both tests found significant mean differences (see below) at the $p = .05$ level between Swedish and German listeners. No significant mean difference was found between the Finnish and Swedish listeners or the Finnish and German listeners.

Analysis of variance was also conducted at the individual voice level, that is the perceived age for each voice was a separate dependent variable. No significant difference in response due to country of residence was found for the Australian female voice ($F_{2,147} = .299$; $p = .742$), or the US male voice ($F_{2,146} = 1.825$; $p = .165$). Significant differences at the $p = .05$ level were found for the remaining four voices: the New Zealand female voice ($F_{2,149} = 4.582$; $p = .012$), the New Zealand male voice ($F_{2,143} = 5.322$; $p = .006$), the US female voice ($F_{2,144} = 4.135$; $p = .018$) and the Australian male voice ($F_{2,143} = 4.051$; $p = .019$). Tukey and Scheffé posthoc tests were applied to the dependent variable perceived age for each of the voices for which the analysis of variance had shown to be significant. Both tests found significant mean differences at the $p = .05$ level between Swedish and Finnish listeners for the Australian male and New Zealand female voices, and between the Swedish and German listeners for the US female voice and the New Zealand and Australian male voices. No significant mean difference between the German and the Finnish listeners was found for any of the four voices.

In both of the cases where a significant difference in means was found between the Finnish and Swedish listeners, the Finnish listeners tended to rate the speakers' ages as younger than the Swedish listeners. They thus made a greater number of incorrect perceived age estimations. In those cases where the German means differed significantly from the Swedish means, there is no general trend. For the Australian male voice the German listeners followed the Finns and assigned younger ages than the Swedes. For the New Zealand male and the US female voice the Germans assigned older ages than the Swedes. In the case of the New Zealand male voice this made their assignments less accurate, whereas in the case of the US female voice this made the Germans more accurate than the Swedish listeners.

An analysis of variance was also conducted to examine the more general question of whether English as a first or foreign language, that is a non-native language in a non-English speaking society, resulted in any significant difference in perceived estimation of ages based on vocal cues alone. Here, native/non-native language was the fixed factor and perceived age was the dependent variable. No significant difference was found ($F_{1,1788} = .005$; $p = .943$).

DISCUSSION

The lack of impact on subjective age assignment due to first language confirms Braun and Cerrato's (1999) finding. Further, the lack of significance of local versus non-local English accent and dubbing versus subtitled policy would appear to rule out any impact on perceived age estimation due to the media. It would appear at first glance that the differences which exist in the data analyzed in this paper

would have to be attributed to the difficulties of assigning age to particular voices as compared to identifying their accent. However, the fact that the American listeners were more accurate at estimating the age of the US voices and less accurate with the ages of the New Zealand female and Australian male voices than the New Zealand and Australian listeners, yet no clear similar effect was found for the New Zealand or Australian voices, hints that media exposure does have some degree of impact and that further research is necessary to provide a more conclusive answer to the question, perhaps by using actual age rather than age group assignment.

REFERENCES

- Bayard, D. (1987) "Class and change in New Zealand English: a summary report", *Te Reo*, 30: 3–36.
- Bayard, D. (2000) "The cultural cringe revisited: changes through time in Kiwi attitudes toward accents", in Bell, A. and Kuiper, K. (Eds) *New Zealand English*, 297–322. Amsterdam/Wellington: John Benjamins/ Victoria University Press.
- Bayard, D. & Sullivan, K.P.H. (2000) "Perception of country of origin and social status of English speakers by Swedish and New Zealand listeners", *Proceedings of FONETIK 2000*, 33–36.
- Bayard, D., Weatherall, A., Gallois, C. & Pittam, J. (2001) "Pax Americana? Accent attitudinal evaluations in New Zealand, Australia, and America", *Journal of Sociolinguistics*, in press.
- Braun, A. & Cerrato, L. (1999) "Estimating speaker age across languages" *Proceedings of ICPHS99*, 1369–1372.
- Cerrato, L., Falcone, M. & Paoloni, A. (1998) "Age estimation of telephonic voices", *Proceedings of the RLA2C*, 20–24.
- Neiman, G.S. & Applegate, J.A. (1990) "Accuracy of listener judgements of perceived age relative to chronological age in adults", *Folia phoniatrica*, 42: 327–320.
- Norris, P., Pauling, B., Lealand, G., Huijser, H. & Hight, C. (1999) *Television in Ten Countries: Local content and diversity*, source: http://www.nzonair.govt.nz/local_content99/index.html
- Rose, P. & Duncan, S. (1995) "Naïve auditory identification and discrimination of similar voice by familiar listeners", *Forensic Linguistics* 2(1): 1–17.
- Schlichting, F. & Sullivan, K.P.H. (1996) "Discrimination of imitated voices" *Proceedings of SST '96*, 103–108.
- Schlichting, F. & Sullivan, K.P.H. (1997) "The imitated voice — a problem for voice line-ups?", *Forensic Linguistics* 4(1): 148–165.
- Sullivan, K.P.H. & Y.N. Karst "Perception of English accent by native British English speakers and Swedish learners of English", *Proceedings of SST '96*, 509–514.
- Sullivan, K.P.H. & Schlichting, F. (2000) "Speaker discrimination in a foreign language: first language environment, second language learners" *Forensic Linguistics* 7(1): 95–111.
- Weatherall, A., Gallois, C. & Pittam, J. (2000) [dated 1998]. "Australasians identifying Australasian Accents", *Te Reo* 41: 153-62.