Available online at www.sciencedirect.com
ScienceDirect
Computer Speech and Language 21 (2007) 641–651
COMPUTER SPEECH AND LANGUAGE
www.elsevier.com/locate/csl

Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners

Caroline Jones [a,*], Lynn Berry [b], Catherine Stevens [c]

[a] *School of Education, University of New South Wales, Sydney 2052, New South Wales, Australia*
[b] *VerifyMe P/L*
[c] *MARCS Auditory Laboratories, University of Western Sydney, Sydney, New South Wales, Australia*

Received 11 August 2006; received in revised form 5 March 2007; accepted 5 March 2007
Available online 16 March 2007

Abstract

This experiment assessed the effect of variation in speech rate on comprehension and persuasiveness of a message presented in text-to-speech (TTS) synthesis to native and non-native listeners. Eighty non-native speakers of English and 80 native speakers of Australian English were randomly assigned to listen to eight banking product descriptions under one of four conditions: normal rate (155 words per minute) with no background noise, normal rate with multi-talker background noise (+6 dB SNR), fast-normal (178 words per minute) with no background noise, and fast-normal with multi-talker background noise. Participants completed comprehension tests and rated each product's usefulness. A faster rate lowers comprehension for both native and non-native listeners but does not influence the persuasiveness of the message. The findings have implications for the selection of speech rates for persuasive messages delivered to native and non-native listeners using TTS.

Keywords: Synthetic speech; Comprehension; Second language; Intelligibility; Persuasion

1. Introduction

Research into intelligibility of degraded natural speech shows that speech rate and level affect intelligibility to different extents for differently aged listeners (e.g., Gordon-Salant and Fitzgibbons, 1993, 1995; Pichora-Fuller et al., 1995; Tun, 1998). By comparison with natural speech, text-to-speech (TTS) can be also regarded as a kind of impoverished speech signal; although TTS systems have improved greatly in recent years most TTS systems are still recognizably different from natural speech due to poorer clarity, odd intonation and rhythm (Bailly, 2002; Koul and Allen, 1993; Pols et al., 1998). The research demonstrating that age of listener is a factor in the intelligibility of degraded speech presented at different rates and levels can therefore be viewed as underlining a need for research into how rate and level affect the experience of TTS for listeners of different

* Corresponding author. Tel.: +61 2 9385 1441; fax: +61 2 9385 1946.
E-mail address: sc.jones@unsw.edu.au (C. Jones).

0885-2308/$ - see front matter
doi:10.1016/j.csl.2007.03.001

ages and different audiometric profiles. To these kinds of listener profiles (age and hearing status), which may affect sensory perception and/or cognitive processing of speech, can be added the effects of a non-native language background. The effect of rate and signal-to-noise ratio on native vs. non-native listeners' experience of TTS is the focus of the present study.

Research into different listeners' needs for TTS is starting to be done with modern TTS systems. In a study by Reynolds and Jefferson (1999) child listeners (6–7 and 9–11 years) had shorter response latencies in a sentence verification task with natural speech than with DECtalk. This may indicate that speech perception, which continues to mature through childhood and teenage years (Johnson, 2000) renders TTS speech, which is acoustically relatively impoverished (Luce et al., 1983) especially challenging for young children to process.

An effect of listeners' language knowledge is evident in a recent study by Axmear et al. (2005). Axmear et al. (2005) compared the intelligibility of sentences produced in DECtalk Perfect Paul for child listeners from either a monolingual English background or balanced bilingual background (English and another language at equal proficiency as rated by parents). Results indicated little reduced intelligibility by bilingual child listeners (at 92% mean accuracy) vs. monolingual child listeners (at 99%) for sentences in natural speech. For TTS, however, there was a large decrement in intelligibility for bilingual children (at 61%) as compared with monolingual children (at 84%).

TTS intelligibility for adult non-native listeners has been compared with that for native adult listeners in just two other studies. Reynolds et al. (1996) found that non-native listeners were less accurate than native listeners in transcribing English sentences presented in DECtalk. This effect was especially strong in noise (10 vs. 0 dB SNR). An earlier study by Mack et al. (1990, cited by Axmear et al., 2005) also found a native listener advantage in TTS: non-native listeners performed more poorly at CVC identification in synthetic speech but not natural speech. The present study builds on this previous research through an investigation of the effects of speech rate and noise level on the experience of TTS speech by native vs. non-native adult listeners. The experience of listeners differing in language background is assessed by two measures: comprehension (a measure of intelligibility) of the TTS message as affected by rate and noise level, and via the persuasiveness as affected by rate.

Previous research into speech rate indicates that for natural speech at rates of 125–225 words per minute, increasing speech rate is non-significantly associated with decreasing recall, for native listeners (Harwood, 1955; Nelson, 1948). Above 225 words per minute, there is an accelerating decline in comprehension by native listeners (Foulke and Sticht, 1969). Based on this previous research, the prediction for comprehension (intelligibility) of TTS is that a faster rate, within the bounds of normal speech, will lead to worse comprehension for non-native but not native listeners. This language background effect is predicted to be greater in the presence of moderate background noise (6 dB SNR). At signal-to-noise ratios of 0 or 5 dB multi-talker babble is known to reduce the intelligibility of TTS, and to reduce it more than for natural speech (Venkatagiri, 2003). Noise is also known to affect non-native listeners more. High proficiency non-native speakers tested by Wijngaarden et al. (2002) required a signal-to-noise ratio for natural speech that was 1–7 dB higher than that of native speakers to obtain 50% sentence intelligibility. TTS transcription accuracy of non-native listeners studied by Reynolds et al. (1996) was more adversely affected by noise (+10 dB SNR babble) than transcription by native listeners.

The interest in the persuasiveness of the TTS message as affected by rate is based on research from speech science and social psychology (e.g., Addington, 1968; Apple et al., 1979; Gunderson and Hopper, 1976; LaBarbera and MacLachlan, 1979; MacLachlan and Siegel, 1980; MacLachlan, 1979, 1982; Mehrabian and Williams, 1969; Miller et al., 1976; Moore et al., 1986; Schlinger et al., 1983; Smith et al., 1975; Stephens, 1982) and human factors and human–computer interaction (e.g., Nass and Lee, 2001; Nass et al., 1997; Stern et al., 1999). Research with a social psychology and commercial orientation has suggested that for American listeners at least, a faster rate will lead the speaker to be regarded as more competent and knowledgeable (e.g., MacLachlan and Siegel, 1980). It seems likely that the relationship between speech characteristics and perceptions of the speaker is likely to involve complex, many-to-one mappings which may also vary with culture. For example, Lee and Boster (1992) showed that Korean males but not females were rated more favourably when they spoke slowly than when they spoke quickly. The mappings between speech characteristics and perceptions of the speaker are likely to be many-to-one because of the likely effect on persuasiveness of other prosodic factors apart from speech rate. Research into prosodic characteristics of different speaking styles, such as storytelling for TTS applications (Theune et al., 2006) makes it clear that as a speaker's communicative intent changes, there are changes in patterns of fundamental frequency, pauses, and amplitude. There are thus likely to be properties other than speech rate which characterise persuasive speech.

As an investigation into the effectiveness of manipulations of speech rate alone in altering persuasiveness for non-American listeners, this study was intended as exploratory. A straightforward extension of ideas and results from American English research with natural speech would predict that a faster rate would lead to greater persuasion. The present research was designed to test whether that would be true for an English speaking but non-American sample, and true at much less extreme values of speech rate manipulation than have been used in most prior investigations of the relationship between speech rate and persuasion.

A further feature of the experimental approach chosen here is the use of an implicit measure of persuasion. An implicit measure of persuasion was chosen over a direct measure to avoid demand characteristics (Orne, 1962); the current study aims to measure how a factor like speech rate actually affects persuasiveness in TTS, rather than what participants might think would affect persuasiveness in TTS. To this end, instead of asking participants to directly rate the perceived persuasiveness of a virtual seller, participants in the current study were asked to rate the perceived usefulness of each product. By comparing how usefulness ratings differ as a function of speech rate, it is possible to infer something about the relationship between speech rate and persuasion, for the specific speech rates and listener groups examined in the study.

2. Method

2.1. Design

Two dependent variables were investigated: listening comprehension (a measure of intelligibility) and perceived product usefulness (an implicit measure of persuasion). The experiment had a 2 (normal vs. fast speech rate) × 2 (noise vs. no noise) × 2 (native vs. non-native listener) factorial design. Rate and noise were manipulated as between-subjects factors, to guard against demand characteristics that would occur when listeners react to the perceived difference between conditions (e.g., as discussed for speech rate in Moore et al., 1986, Experiment 2).

The TTS system was a commercially available Australian English male voice in its demonstration version available on the internet in April 2004. This is a fairly high-quality TTS system for its time; to native listeners it sounds more Australian than American or British, there are only very few detectable segmental errors in the stimulus audio texts generated with the demonstration version of this TTS system and, although the intonation and rhythm do not sound entirely natural, it is not monotonic or overly machine-like.

2.2. Participants

Participants were 160 adults recruited via posters and emails from the university student and staff populations at University of Western Sydney (Bankstown campus) and University of New South Wales, in Sydney, Australia. Eighty participants were native speakers of Australian English, i.e. people who had been born and grown up in Australia speaking English. Some were bilingual in English and another language or languages. All native participants were students recruited from University of Western Sydney and received research credit towards their undergraduate psychology course.

Non-native participants were required to have arrived in Australia after the age of 15 years, and to have grown up speaking a language or languages other than English. The cutoff of 15 years old at age of arrival was employed to sample later second language (L2) learners, rather than earlier L2 learners who tend to have relatively less difficulty with spoken English syntax and phonology (Flege et al., 1999) whether or not there are maturational constraints on second language learning as argued by, for example, Johnson et al. (1996). The 80 non-native participants included undergraduate psychology students (participating for research credit) and some graduate psychology students and staff (volunteers) from the University of Western Sydney ($n = 44$), as well as undergraduate and graduate students and some staff (participating for $10) from the University of New South Wales ($n = 36$). Volunteer and paid participants were distributed equally across the four (rate × noise) conditions. All participants were naive to the purpose of the study. None of the participants were students or researchers in speech science or speech technology and it was therefore assumed that they would represent the range of experience with synthetic speech in the general population. Non-native participants did not differ significantly in age, length of residence in Australia, or age of arrival, across conditions. Overall, 68% of the non-native listeners were women. The mean age of non-native participants was 28.2 years (s.d. 8.4, min. 18, max. 57). Their mean age of arrival in Australia was 22.9 years (s.d. 6.2, min. 12, max. 44). Their mean length of residence in Australia was 5.3 years (s.d. 6.5, min. 0.2, max. 36). Within each group, there was a considerable range of native languages, but approximately similar proportions across groups of different area language backgrounds. The majority of participants in each group were speakers of a language of South-East or East Asian language (self-identified language was Mandarin, Cantonese, Chinese, Japanese, Vietnamese, Indonesian or Thai). In each group there were also 2–4 speakers of a Western European language (French, German or Italian) which is relatively more similar to English and young speakers of these languages have often studied written and spoken English for several years at school. Each group also contained 1–3 speakers of languages of the Indian subcontinent and/or the Pacific such as Hindi, Telugu, Sinhalese and Filipino. Each group also contained 2–6 speakers of other languages such as forms of Arabic, or Eastern European languages such as Russian, Ukrainian, and Bulgarian. The heterogeneity of the non-native listener group in terms of language background was intentional; it reflects the mix of languages spoken in Sydney, a multicultural city where 33% of residents were born overseas and 30% speak a language other than English (Australian Bureau of Statistics, 2001).

2.3. Stimuli

The stimuli were descriptions of eight banking products (see Appendix A), presented in audio form using a demonstration version of an Australian English male TTS synthesis system available on the internet in April 2004. The eight product descriptions were fictional, but resemble offers of financial software and financial planning tools widely advertised on the internet at the time of this study (e.g., interest rate calculator, identity theft package, home insurance tool). Pilot testing had led to the selection of the final eight products from a pool of twelve product descriptions as ones which had a similar, moderate level of appeal to university student and staff participants. Each product description was quite short, with a mean of 47 words (s.d. 1.6 words). In audio form, the mean duration of each product description was 18 s (s.d. 1.0) in the 155 wpm condition, and 16 s (s.d. 0.9) in the 178 wpm condition. Participants rated each product on a scale of 1–5 (where 1 = not useful at all, and 5 = extremely useful), and answered two true–false questions about the product description they had just heard (e.g., This software allows you to compare your expenses to normal spending patterns. TRUE/FALSE. See Appendix B for the rating question and the true–false questions).

Both rate and noise manipulations are well within normal limits of variation in English. The normal speech rate employed was 155 words per minute (wpm), or 237 syllables per minute (spm), which is approximately equal to the mean Australian English speech rate found by Block and Killen (1996). The faster speech rate employed in the experiment is 178 wpm, or 272 spm. Block and Killen (1996) measured speech rate for 60 adults aged 21–30 years, from a variety of geographical and socioeconomic regions within Melbourne. Mean speech rates were similar in reading and conversation, slightly slower for reading (mean 230 spm, s.d. 28) than conversation (mean 238, s.d. 21), with no gender differences in either condition. The range (minimum to maximum) in reading was 181–281 spm for females, and 171–287 spm for males. In conversation the range was 198–282 spm for females, and 196–277 spm for males. The only other study of speech rate in Australian English is Bernard (1965) who measured speech of 16–18 year old males, and found a mean speech rate of 222 spm.

Speech rate variation was achieved by first generating audio text using the TTS system at its demonstration speed (178 wpm) and then resynthesizing the audio text at the slower rate, i.e. at approximately an average Australian English speech rate, using a linear lengthening function within Praat version 4.2.31 (Boersma and Weenink, 2004). The originally generated 178 wpm audio text was also resynthesized at 100% of its rate, so that in both rate conditions the audio stimuli were resynthesized speech. All audio recordings were normalized to 95% maximum amplitude using CoolEdit2000.

Continuous multi-talker babble was added to copies of the resulting TTS stimuli to produce stimuli for the noise condition. The multi-talker babble was added at a level which resulted in a mean signal-to-noise ratio of 6 dB in the noise condition. This signal-to noise ratio was selected on the basis of previous research (Stevens et al., 2005; Venkatagiri, 2003; Wijngaarden et al., 2002) and subjective assessment by native speaker experimenters as a distracting but realistic noise level for background conversation.

2.4. Procedure

Testing was conducted mostly in small groups of 2–5 participants. Owing to test scheduling, some participants were tested individually. The audio stimuli were played from a pre-recorded CD on a laptop, through high-quality powered speakers. All participants were seated at a distance of about 2–3 m from the audio speakers, which were set at a constant volume, approximately 75 dB SPL. Each audio description was played once. To minimize the effect of any one product on participant responses, each participant was randomly assigned to one of four different presentation orders of the eight banking products. After hearing each audio condition, participants turned the page, rated the product's usefulness and then answered two true–false questions. A period of 30 s was provided for the ratings and these answers. Then a warning tone sounded, followed after 10 s by the next audio description. Testing took approximately 8 min in total.

2.5. Analysis

A mean comprehension score (proportion correct) was calculated for each participant based on 13 true–false questions. Of the 16 original true–false questions (see Appendix B) three questions were not included in the analysis (2a, 3b, 6a), as it was realised after the experiment that they asked participants to judge the truth/falsity of product information that was not actually provided in the description. These questions were not included in the analysis as answers to them were more likely to reflect memory for the items rather than comprehension. A mean rating score (out of 5) was also calculated for each participant based on the ratings given to the eight products.

3. Results

3.1. Comprehension accuracy results

Mean comprehension accuracy (with error bars showing standard error of the mean) is shown as a function of speech rate and background noise level, for native listeners in Fig. 1 and for non-native listeners in Fig. 2. The accuracy data were analysed using a $2 \times 2 \times 2$ ANOVA with language background, speech rate and noise as between-subjects factors.

In comprehension accuracy, there is no main effect of Noise, $F(1,152) = 1.22$, $p = 0.27$. There is also no main effect of Language Background, $F(1,152) = 0.48$, $p = 0.49$. There is, however, an effect of rate:

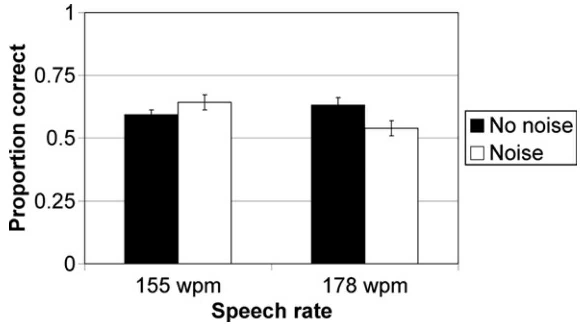

Fig. 1. Mean comprehension accuracy for native English listeners, by speech rate and noise condition. (Error bars represent standard errors of the mean.)

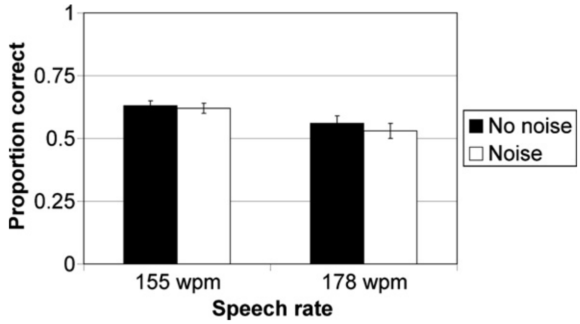

Fig. 2. Mean comprehension accuracy for non-native English listeners, by speech rate and noise condition. (Error bars represent standard errors of the mean.)

comprehension accuracy is lower at the faster speech rate (178 wpm) than the slower rate (155 wpm), $F(1, 152) = 7.89$, $p = 0.006$. The rate by noise interaction approaches significance, $F(1, 152) = 3.53$, $p = 0.06$. (Judging from Figs. 1 and 2, the tendency is for comprehension to be even worse at 178 wpm than 155 wpm when there is more background noise.) There is no significant three-way interaction between rate, noise and language background, $F(1, 152) = 3.03$, $p = 0.08$, and the remaining two-way interactions are also non-significant: rate by language background ($F(1, 152) = 1.31$, $p = 0.25$), and noise by language background ($F(1, 152) = 0.001$, $p = 0.98$).

As the comprehension scores are generally not high, the question whether the scores are above chance is of interest. For native listeners, when comprehension scores are assessed against chance (50%, in true/false questions), performance is above chance in the 155 wpm rate conditions (95% C.I. for No Noise = 0.53–0.65, for Noise = 0.58–0.70) and in the 178 wpm no noise condition (95% C.I. = 0.57–0.69). Performance is not different from chance in the 178 wpm Noise condition (95% C.I. = 0.48–0.60).

For non-native listeners, in comparisons against chance, performance is above chance for both 155 wpm conditions (95% C.I. no noise = 0.58–0.69; for Noise = 0.56–0.67) and for the 178 wpm no noise condition (95% C.I. = 0.51–0.61). Performance is not different from chance in the 178 wpm noise condition (95% C.I. = 0.48–0.59).

3.2. Product usefulness rating results: implicit measure of persuasion

Mean usefulness rating (with error bars showing standard error of the mean) – an implicit index of persuasion – is shown as a function of speech rate and background noise in Fig. 3 for native listeners and Fig. 4 for non-native listeners. Usefulness ratings were analysed using a $2 \times 2 \times 2$ analysis of variance (ANOVA) with

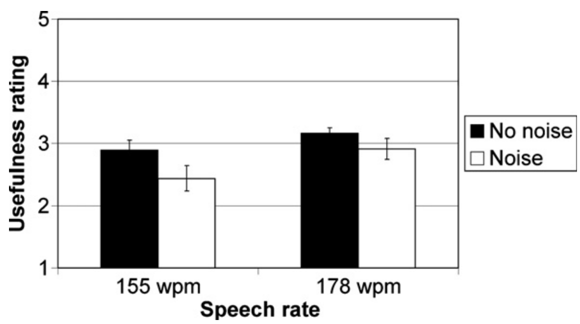

Fig. 3. Mean product usefulness rating for native English listeners, by speech rate and noise condition. (Error bars represent standard errors of the mean.)

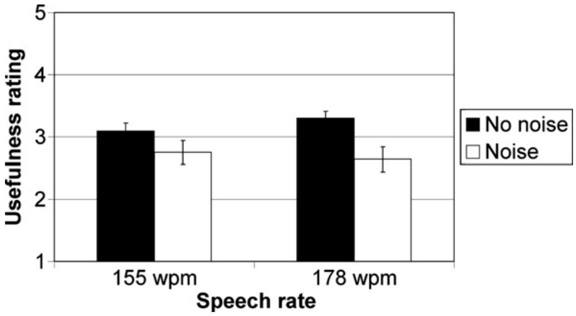

Fig. 4. Mean product usefulness rating for non-native English listeners, by speech rate and noise condition. (Error bars represent standard errors of the mean.)

between-subjects factors of language background (native in English vs. non-native in English), speech rate (155 wpm vs. 178 wpm) and noise (noise vs. no noise).

Noise is the only factor to have a significant effect on rated product usefulness. Participants gave higher product usefulness ratings in the no noise condition (mean = 3.1; 95% C.I. = 3.0–3.3) than the noise condition (mean = 2.7; 95% C.I. = 2.5–2.8), $F(1,152) = 14.30$, $p < 0.001$. Rated product usefulness did not vary by listener's language background or by speech rate. There was a non-significant tendency, however, for higher product usefulness ratings where the product description was delivered at the faster speech rate (178 wpm vs. 155 wpm), $F(1,152) = 3.33$, $p = 0.07$. There was no interaction of rate with noise, $F(1,152) = 0.06$, $p = 0.80$. There was no effect of language background on rated product usefulness, $F(1,152) = 0.73$, $p = 0.40$, and no interaction of language background with rate ($F(1,152) = 1.98$, $p = 0.16$) or noise ($F(1,152) = 0.44$, $p = 0.51$). The three-way interaction between language background, rate and noise was also non-significant, $F(1,152) = 1.35$, $p = 0.25$.

4. Discussion

The predictions for comprehension (intelligibility) were that at the faster speech rate as compared with the slower speech rate, comprehension would be worse for the non-native listeners but not for the native listeners. This interaction between speech rate and language background was also predicted to be greater in the noise condition. The results for comprehension did not conform to these expectations as there was no interaction between rate and language background. Instead, there was an overall effect of rate; comprehension was lower at the faster speech rate for both native and non-native listeners. The direction of this effect of rate was as predicted for the non-native listeners but extended surprisingly to the native listeners as well.

The prediction for persuasion was that the TTS message would be more persuasive at the faster speech rate. This prediction was not borne out by the data. There was no difference between native and non-native listeners in any effect of rate on persuasion. The presence of background noise, for which there was no expected effect, proved to reduce the persuasiveness of the TTS message for both native and non-native listeners.

These results have practical implications for the implementation of TTS systems. In the present study, comprehension of message delivered by TTS is worse at the faster (178 wpm) speech rate for both native and non-native listeners. This is somewhat surprising as the faster speech rate is also the default speech rate for the demonstration version of this TTS system and lies well within the range of normal speaking rates in Australian English (and other varieties of English). The speech rate for TTS which resulted in better comprehension was 155 wpm. This rate is known from descriptive studies of speech production to be more similar to normal rates of spoken English in Australia. This suggests that for spoken messages with the same level of complexity as those in the current study, such as financial advertisements and other information, Australian listeners understand the message better if the speech rate is close to the normal rate of spoken Australian English. Implementation of TTS systems may well be served by paying close attention to what is becoming known about normal speech rates in particular populations. These are the rates for which listeners have likely adapted their perceptual processing.

The implications of the present results for the use of modern TTS in noise are relatively encouraging, given that TTS systems must sometimes be used in the presence of everyday background noise (such as other conversations, traffic, etc.). Although comprehension levels are overall not high in this study, the presence of moderate background babble noise at the level manipulated in this study (6 dB SNR) appears to have less of a detrimental effect on comprehension than increasing speech rate to 178 wpm.

For TTS implementations in which persuasiveness is desirable, the present study indicates that simply raising the speech rate from normal rate to the high end of normal rate does not result in greater persuasiveness. There is no benefit to be had from this manipulation alone, for a target audience of native or non-native listeners. As raising the speech rate above 178 wpm is likely to reduce comprehension even further for all listeners, it would not appear that for most applications an increase in speech rate is worth pursuing to increase persuasiveness. Progress in this area is most likely to benefit from in-depth descriptive studies of the acoustic characteristics of human speakers who are speaking with persuasive intent, possibly speakers of different genders, with different personalities, and in different cultures and languages. As the results of such studies are implemented into TTS applications and these applications are evaluated, we are likely to learn more about the required characteristics of persuasive speech for TTS systems.

Acknowledgements

The authors wish to thank Kristy Gatley, Kimberley Hastie, and Carlie Hosking for their assistance in recruiting participants and collecting some of the data.

Appendix A. Stimulus texts (banking product descriptions)

1. This tax software guides you through your annual return step by step. It asks you simple personalized questions and it outputs your answers into the correct cells on your tax return. The software reflects the latest tax law reform, and double-checks the maths for accuracy.
2. Superannuation choices can be hard to make. Although we all look forward to claiming our Super when we retire, very few people actually understand how Super works. Find out how much you will be able to claim, and plan now for your future using superannuation software.
3. An accident in the workplace, on the roads, or at home can leave you unable to do your job. Disablement can impact on your finances whether the disablement is temporary or permanent. Use this manager to determine whether your current cover is appropriate for your age, gender, and occupation.
4. Everyone has slightly different spending habits although there are predictable expenses, such as the weekly shopping trip and your rent or mortgage. Where do you spend your money? This budget manager keeps a record of your expenses and allows you to compare that to your ideal spending pattern.
5. Did you know that insurance companies determine your rates based in part on your property claims history? Whether you rent a property or own your own home it is important to know your home insurance score. Improve your chances for lower premiums with this easy to use product.
6. Identity theft can ruin your reputation and claim your savings before you even know it. The identity theft package helps protect you from identity theft, and minimizes the impact if you do become a victim. It is a small price to pay for peace of mind.
7. The interest rate calculator is an easy-to-use mortgage calculator. You can plan for your first home or calculate the best way to pay off your mortgage in the shortest time. The calculator includes a monthly payment chart, an annual table, and a flexible graphics function.
8. This credit planning tool saves time and effort when planning major expenditures. The comprehensive credit report contains credit information you need from your bank, your employer, and the Australian Taxation Office combined in a single, easy-to-read report. So when are you getting that special present for yourself?

Appendix B. Rating item and true–false questions

Rating item

How useful does this product seem to you? (Please circle one number on the scale).

|----------|----------|----------|----------|
1 2 3 4 5

Not useful at all Very useful

B.1. True–false questions

1.
a. Everyone has some predictable kinds of expenses, such as rent or mortgage and weekly shopping.
TRUE/FALSE (please circle one).
b. This software allows you to compare your expenses to normal spending patterns.
TRUE/FALSE (please circle one).
2.
a. Disablement impacts more on your finances when the disablement is permanent rather than temporary.
TRUE/FALSE (please circle one).
b. You can use this calculator to see if your cover is appropriate for your occupation, gender, and age.
TRUE/FALSE (please circle one).
3.
a. Insurance companies calculate your premium based mostly on your property claims history.
TRUE/FALSE (please circle one).
b. Home owners are more likely than renters to know their home insurance point score.
TRUE/FALSE (please circle one).
4.
a. You can lose your savings and your good reputation through identity theft.
TRUE/FALSE (please circle one).
b. The identity theft package helps you avoid identity theft and protects you fully against any loss due to identity theft.
TRUE/FALSE (please circle one).
5.
a. The interest rate calculator is easy to use and you can use it to plan your mortgage.
TRUE/FALSE (please circle one).
b. The interest rate calculator is useful for people who are thinking of buying a first home and also for people who are paying off their mortgage.
TRUE/FALSE (please circle one).
6.
a. Young people often find superannuation choices hard to make.
TRUE/FALSE (please circle one).
b. The superannuation software helps you calculate how much money you will have at retirement age.
TRUE/FALSE (please circle one).
7.
a. This tax software is designed to help you fill out an annual tax return.
TRUE/FALSE (please circle one).
b. The software asks everyone the same simple questions and outputs your answers into the correct cell on your tax return.
TRUE/FALSE (please circle one).

8.
a. This tool helps you organise your credit card expenses.
TRUE/FALSE (please circle one).
b. This tool helps you plan purchases by combining financial information from three different sources.
TRUE/FALSE (please circle one).

References

Addington, D., 1968. The relationship of selected vocal characteristics to personality perception. Speech Monographs 35, 492–503.

Apple, W., Streeter, L., Krauss, R., 1979. Effects of pitch and speech rate on personal attributions. Journal of Personality and Social Psychology 37, 715–727.

Australian Bureau of Statistics, 2001. 2001 Census of Population and Housing. Australian Government, Canberra.

Axmear, E., Reichle, J., Alamsaputra, M., Kohnert, K., Drager, K., Sellnow, K., 2005. Synthesized speech intelligibility in sentences: a comparison of monolingual English-speaking and bilingual children. Language, Speech, and Hearing Services in Schools 36, 244–250.

Bailly, G., 2002. Close shadowing natural versus synthetic speech. International Journal of Speech Technology 6, 11–19.

Bernard, J., 1965. Rates of utterance in Australian dialectic groups. Occasional Paper No. 7. Australian Language Research Centre, The University of Sydney.

Block, S., Killen, D., 1996. Speech rates of Australian English-speaking children and adults. Australian Journal of Human Communication Disorders 24, 39–44.

Boersma, P., Weenink, D., 2004. Praat: doing phonetics by computer (Version 4.2.31) [Computer program]. <http://www.praat.org/> (Retrieved April 30).

Flege, J., Yeni-Komshian, G., Liu, S., 1999. Age constraints on second-language acquisition. Journal of Memory and Language 41, 78–104.

Foulke, E., Sticht, T., 1969. Review of research on the intelligibility and comprehension of accelerated speech. Psychological Bulletin 72, 50–62.

Gordon-Salant, S., Fitzgibbons, P., 1993. Temporal factors and speech recognition performance in young and elderly listeners. Journal of Speech and Hearing Research 36, 1276–1285.

Gordon-Salant, S., Fitzgibbons, P., 1995. Recognition of multiply degraded speech by young and elderly listeners. Journal of Speech and Hearing Research 38, 1150–1156.

Gunderson, D., Hopper, R., 1976. Relationship between speech delivery and speech effectiveness. Communication Monographs 43, 158–165.

Harwood, K., 1955. Listenability and rate of presentation. Speech Monographs 22, 57–59.

Johnson, C., 2000. Children's phoneme identification in reverberation and noise. Journal of Speech, Language, and Hearing Research 43, 144–157.

Johnson, J., Shenkman, K., Newport, E., Medin, D., 1996. Indeterminacy in the grammar of adult language learners. Journal of Memory and Language 35, 335–352.

Koul, R., Allen, G., 1993. Segmental intelligibility and speech interference thresholds of high quality synthetic speech in presence of noise. Journal of Speech and Hearing Research 36, 790–798.

LaBarbera, P., MacLachlan, J., 1979. Time-compressed speech in radio advertising. Journal of Marketing 43, 30–36.

Lee, H., Boster, F., 1992. Collectivism–individualism in perceptions of speech rate. Journal of Cross-Cultural Psychology 23, 377–388.

Luce, P., Feustel, T., Pisoni, D., 1983. Capacity demands in short-term memory for synthetic and natural speech. Human Factors 25, 17–32.

MacLachlan, J., 1979. What people really think of fast talkers. Psychology Today 13, 113–117.

MacLachlan, J., 1982. Listener perception of time-compressed spokespersons. Journal of Advertising Research 2, 47–51.

MacLachlan, J., Siegel, M., 1980. Reducing the costs of television commercials by use of time compression. Journal of Marketing Research 17, 52–57.

Mehrabian, A., Williams, M., 1969. Nonverbal concomitants of perceived and intended persuasiveness. Journal of Personality and Social Psychology 13, 37–58.

Miller, N., Maruyama, G., Beaber, R., Valone, K., 1976. Speech of speech and persuasion. Journal of Personality and Social Psychology 34, 615–624.

Moore, D., Hausknecht, D., Thamodaran, K., 1986. Time compression, response opportunity, and persuasion. The Journal of Consumer Research 13, 85–99.

Nass, C., Lee, K., 2001. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. Journal of Experimental Psychology (Applied) 7, 171–181.

Nass, C., Moon, Y., Green, N., 1997. Are machines gender neutral? Gender-stereotypic responses to computers with voices. Journal of Applied Social Psychology 27, 864–876.

Nelson, H., 1948. The effect of variation of rate on the recall by radio listeners of "straight" newscasts. Speech Monographs 15, 173–180.

Orne, M., 1962. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. American Psychologist 17, 776–783.

Pichora-Fuller, M., Scheider, B., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. Journal of the Acoustical Society of America 97, 593–608.

Pols, L., van Santen, J., Abe, M., Kahn, D., Keller, E., 1998. The use of large text corpora for evaluating text-to-speech systems. In: Proceedings of the First International Conference on Language Resources and Evaluation. Granada, Spain.

Reynolds, M., Jefferson, M., 1999. Natural and synthetic speech comprehension: comparison of children from two age groups. Augmentative and Alternative Communication 15, 174–182.

Reynolds, M., Bond, Z., Fucci, D., 1996. Synthetic speech intelligibility: comparison of native and non-native speakers of English. Augmentative and Alternative Communication 12, 32–36.

Schlinger, M., Alwitt, L., McCarthy, K., Green, L., 1983. Effects of time compression on attitudes and information processing. Journal of Marketing 47, 79–85.

Smith, B., Brown, B., Strong, W., Rencher, A., 1975. Effects of speech rate on personality perception. Language and Speech 18, 145–152.

Stephens, N., 1982. The effectiveness of time-compressed television advertisements with older adults. Journal of Advertising 11, 48–76.

Stern, S., Mullennix, J., Dyson, C., Wilson, S., 1999. The persuasiveness of synthetic speech versus human speech. Human Factors 41, 588–595.

Stevens, C., Lees, N., Vonwiller, J., Burnham, D., 2005. On-line experimental methods to evaluate text-to-speech (TTS) synthesis: effects of voice gender and signal quality on intelligibility, naturalness and preference. Computer Speech and Language 19, 129–146.

Theune, M., Meijs, K., Heylen, D., Ordelman, R., 2006. Generating expressive speech for storytelling applications. IEEE Transactions on Audio, Speech, and Language Processing 14, 1137–1144.

Tun, P., 1998. Fast noisy speech: age differences in processing rapid speech with background noise. Psychology and Aging 13, 424–434.

Venkatagiri, H., 2003. Segmental intelligibility of four currently used text-to-speech synthesis methods. Journal of the Acoustical Society of America 113, 2095–2104.

Wijngaarden, S., Steeneken, H., Houtgast, T., 2002. Quantifying the intelligibility of speech in noise for non-native listeners. Journal of the Acoustical Society of America 111, 1906–1916.