EFFECTS OF INTRA-TALKER DIFFERENCES ON SPEECH UNDERSTANDING IN NOISE BY YOUNGER AND OLDER ADULTS

Huiwen Goy, M. Kathleen Pichora-Fuller and Pascal van Lieshout
Department of Psychology, University of Toronto at Mississauga, 3359 Mississauga Rd N, Mississauga, Ontario L5L 1C6

ABSTRACT

The purpose of the study was to investigate the effects on intelligibility when listening to a talker who produced speech while being exposed to different types and levels of noise. In addition, we also investigated how the word identification performance of younger and older listeners for these conditions was affected in different listening environments. A male young adult talker recorded the SPIN-R test while listening to different types and levels of noise and following different speaking instructions. Younger and older adults with clinically normal hearing from 250 to 3000 Hz were tested on these recordings in three signal-to-noise ratios. Significant word recognition differences were found between the standard talking condition and loud speech recorded in quiet, despite the fact that all sentences were equated for overall intensity. Older listeners found speech produced in noise beneficial when there was more background noise, but younger listeners did not. Clear speech in this study did not produce any benefit for listeners relative to the baseline condition. Acoustic analyses showed that intensity fluctuations within the sentences led to a higher intensity for target words in the loud speech condition relative to all other conditions, while target words in clear speech had a lower fundamental frequency compared to other conditions. Listener performance may change on a test of speech intelligibility when speech is produced under more ecologically valid conditions; however, these effects are small and may be more apparent in older adults and when task difficulty is greater due to lack of contextual support and higher levels of background noise.

SOMMAIRE

Le but de cette étude était d'explorer l'effet sur l'intelligibilité de la parole lorsque celle-ci est produite en présence de différents niveaux et types de bruit. L'effet de ces conditions de production de la parole en présence de différents environnements d'écoute sur les performances à l'identification des mots a été examiné auprés de jeunes adultes et d'adultes plus âgés, présentant des seuils auditifs normaux entre 250 et 3000 Hz. Les phrases du test SPIN-R émises par un locuteur d'âge adulte ont été enregistrées pendant qu'il écoutait différents bruits à différents niveaux et suivait diverses instructions. Ces phrases ont été présentées à trois rapports signal-sur-bruit auprès des deux groupes de participants. Une différence significative a été notée entre le nombre de mots reconnus dans la condition de discours standard et la condition de discours enregistré à volume élévé dans un environnement de silence et ce, même en ayant normalisé le niveau de présentation de toutes les phrases. Les auditeurs plus âgés ont trouvé que le discours produit dans le bruit était plus facile à reconnaître lorsque les phrases étaient présentées avec le bruit de fond au niveau plus haut, mais ce n'était pas le cas pour les jeunes auditeurs. Pour tous les participants, le discours clairement prononcé n'a pas entraîné des performances significativement meilleures, par rapport au discours standard. Les analyses acoustiques ont montré que l'intensité était 2 dB plus haute sur les mots cibles des phrases produites en présence de bruit par rapport aux autres conditions d'enregistrement, alors que la fréquence fondamentale des mots cibles du discours clairement prononcé était plus basse que celle des autres conditions. Les performances aux mesures de reconnaissance de la parole peuvent varier en fonction des conditions dans lesquelles cette parole est produite. Ces effets sont peut-être minimes chez les jeunes adultes, mais peuvent être encore plus apparents chez les personnes âgées et lorsque la tâche de reconnaissance est plus difficile en raison du manque d'indices contextuels et des niveaux élevés de bruit de fond.

1. INTRODUCTION

The success of speech communication depends not only on factors related to the listener, but also on factors related to the talker. It is well-known that speech understanding in noise is more challenging for older listeners than for younger listeners, even when older listeners have good audiograms (Dubno, Dirks, & Morgan, 1982). However, talkers may adjust their speech in an attempt to improve communication when faced with a listener in a difficult listening situation (Smiljanić & Bradlow, 2009). Experiments have demonstrated that changes in speech production due to the talking environment may be beneficial to listeners. For example, speech produced in a noisy environment is more intelligible than speech produced in quiet, when speech is presented to listeners in a noisy environment (Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988; Pittman & Wiley, 2001). Such changes in speech may include an increase in intensity and fundamental frequency (F_0; Letowski, Frank, & Caravella, 1993) as well as changes in articulation (Forrest, Abbas, & Zimmermann, 1986).

Little is known, however, about how younger and older listeners may differ in their word recognition performance when production of speech has been influenced by the environment. The few studies that have compared younger and older listeners have usually tested older adults with hearing loss (Picheny, Durlach, & Braida, 1985; Schum, 1996), and older listeners with hearing loss do not always benefit from the acoustic information available to younger listeners (Ferguson & Kewley-Port, 2002). It is possible that normal-hearing older listeners may benefit more from additional cues compared to younger adults, much as they benefit more from supportive semantic context (Pichora-Fuller, 2008). Conversely, it is also possible that older adults will not benefit from additional acoustic information due to auditory processing deficits; e.g., one study showed that older adults were not able to use voice cues to reduce informational masking (Huang, Xu, Wu, & Li, 2010).

In this study, we investigated three main issues: 1) whether the word recognition performance of listeners would be affected when they heard speech produced in talking conditions other than the typical quiet conditions used for recording stimuli for speech intelligibility tests, 2) whether younger and older listeners would be differentially affected by these changes in talking conditions, and 3) which acoustic changes might underlie any changes in listener performance. We recorded a talker producing speech in different noise environments and under different speaking instructions, and presented those recordings as test sentences to listeners in various levels of background noise.

2. EXPERIMENT 1

2.1 Method

2.1.1 Stimuli recording

The stimuli used were the eight equivalent sentence lists from the Revised Speech Perception in Noise Test (SPIN-R; Bilger et al., 1984). Each list consists of 25 high-context and 25 low-context sentences. High-context sentences contain a sentence-final target word that is highly predictable from the preceding phrase (e.g., Unlock the door and turn the *knob*), while low-context sentences contain an unpredictable sentence-final target word (e.g., We spoke about the *knob*).

For the present study, a new talker was selected to match the original SPIN-R talker as closely as possible on average speaking fundamental frequency (F_0) and speaking rate, in an effort to create a version of the test that was similar to the original except for changes in speech due to the talking environment or speaking instructions. The new talker was selected from a group of six young adult male talkers who spoke Canadian English as their first language. Candidate talkers heard the original SPIN-R sentences in List 1 presented at a level of 70 dB SPL (50 dB HL), using one loudspeaker placed at 225°; this level was chosen as the standard protocol is to present SPIN-R sentences at 50 dB SL (Bilger et al., 1984). Candidates were then asked to repeat each sentence after the talker in their normal voice, using the recordings' speaking rate and intonation as a guide when producing the sentences. An acoustic analysis of each candidate's sentences showed that talker DF was the closest match to the original talker in his average speaking F_0 and speaking rate (121.7 Hz and 4.3 syllables/sec, respectively, compared to 120.5 Hz and 4.3 syllables/sec for the original talker; Kalikow, Stevens, & Elliott, 1977).

Talker DF recorded SPIN-R sentence lists 2 through 8 in different talking environments. The recordings were made while he was seated in a single-walled sound-attenuating International Acoustics Company (IAC) booth, with a Sennheiser Linear E825S microphone placed 18 cm from his lips. His speech was recorded using a Tucker-Davis Technologies System III and the Avaaz Time-Frequency Representation program running on a Dell Precision 360 computer. During the recording sessions, a hardcopy transcript of the sentences was made available to DF, while sentences were presented binaurally through Sennheiser HD265 headphones. All lists were recorded by DF while he wore headphones, so that any speech production changes due to occlusion of the ears were held constant across all conditions. After each sentence was presented, there was a two-second delay before a visual cue appeared on a computer screen to prompt DF to speak, and the rate of presentation of sound files was under the experimenter's control.

The speaking conditions used in this study are summarized in Table 1. SPIN-R lists 2, 7 and 8 were spoken in a quiet talking environment. For List 2, DF was instructed to repeat each sentence using the original talker's speaking rate and intonation (baseline condition). In List 7, he was instructed to "speak clearly" (clear speech condition), and in List 8, to "speak loudly" (loud speech condition). These two speaking styles were elicited for comparison with speech produced in a noisy environment, which may change in both articulation and intensity. Lists 3 and 4 were produced while DF listened to multi-talker babble from the original SPIN-R test, presented at 62 or 66 dB SPL, and Lists 5 and 6 were produced while he listened to speech spectrum noise matched to the babble from the original SPIN-R test, presented at 62 or 66 dB SPL. No specific speaking instructions were given in the noise conditions. Assuming that typical conversational level is about 70 dB SPL, the noise level of 62 dB SPL would match the +8 dB signal-to-noise ratio (SNR) used in the standard protocol for the

SPIN-R test (Bilger et al., 1984). The higher noise level of 66 dB SPL was chosen to create a more challenging condition which is still typical of everyday listening environments (Pearsons, Bennett, & Fidell, 1977). For lists that were spoken in noise, each SPIN-R sentence was presented to DF in quiet, followed by two seconds of noise. The noise continued while a visual cue was presented on a computer screen to prompt him to speak.

Table 1. Summary of speaking conditions (Babble = babble noise; SSN = speech spectrum noise).

SPIN-R List	Recording environment	Speaking instructions
2	Quiet	Follow original talker
7		Speak clearly
8		Speak loudly
3	Babble (62 dB SPL)	None
4	Babble (66 dB SPL)	
5	SSN (62 dB SPL)	
6	SSN (66 dB SPL)	

The average RMS energy of each sentence was equated to 0.05 Pa using a custom MATLAB program. Therefore, the overall sentence intensity level was controlled during the test, but local intensity fluctuations were preserved and could differ across sentences in different talking conditions. The new sentences were aligned with the original SPIN-R babble background, and a 0.6 sec 500-Hz warning tone preceded the presentation of each sentence by 1.25 sec.

2.1.2 Participants

Listeners were 16 younger adults who were undergraduate students from the university (mean age = 19.9 years, SD = 1.7) and 16 community-dwelling older adults who were recruited to participate in studies on healthy aging (mean age = 69.0 years, SD = 4.1). All listeners were native English speakers with pure-tone audiometric thresholds ≤ 25 dB HL from 250 to 3000 Hz in the test ear (Figure 1). None of the participants had heard the SPIN-R test previously. Participants all gave informed consent in compliance with the protocol approved by the institutional ethics review board, and were paid at an hourly rate.

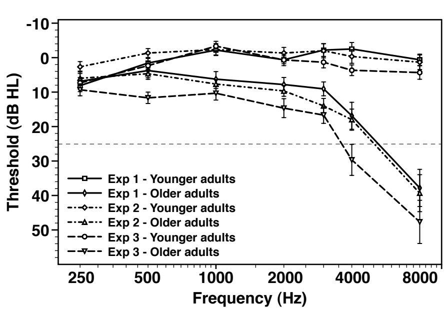

Figure 1. Average pure-tone audiometric thresholds of younger and older listeners in three experiments. Error bars are standard error of the mean. The dotted line is considered to be the limit of clinically normal hearing.

2.1.3 Procedure

Listeners were tested on SPIN-R Lists 2 to 8 by talker DF. They also heard List 1 by the original talker for comparison with DF's baseline condition, but we did not include this list in the current analysis as we were interested in examining intra-talker differences in speech production rather than inter-talker differences. Participants listened to the eight sentence lists in two 1-hour sessions while seated in a double-walled sound-attenuating IAC booth. Sentences were presented monaurally over TDH-50P earphones at 70 dB above the participant's average pure-tone audiometric thresholds at 0.5, 1, and 2 kHz, mixed with SPIN-R babble at 0 dB SNR. Participants were instructed to report the last word of each sentence, and guessing was encouraged. There was no time limit on responding. Responses were scored by the experimenter as they were made, and participants' answers were audio-taped to enable later confirmation of the scoring. The order of talking conditions was counterbalanced across participants.

2.1.4 Data analysis

The data were transformed to satisfy the assumptions of the linear modelling procedure mentioned below. Scores were transformed from raw scores to rationalized arcsine units (RAU) using equations from Sherbecoe and Studebaker (2004)[1].

We modelled listener performance in terms of RAU scores as a function of age group, talking condition, context condition and the interactions between these factors, using a random intercept model with a compound symmetry covariance structure.

2.2 Results

As shown in Figure 2, there was a significant main effect of age, with younger adults obtaining higher scores than older adults, $F(1, 30) = 14.84$, $p < .001$. There was also a significant main effect of context, with higher scores obtained for high-context sentences than for low-context sentences, $F(1, 390) = 1249.60$, $p < .001$, and a significant main effect of talking condition, $F(6, 390) = 13.18$, $p < .001$. Multiple *t*-tests with a Bonferroni correction showed that more words were correctly identified when speech was produced loudly (88%) or in the higher level of babble noise (87%) than when speech was produced in the baseline condition (83%). However, word recognition accuracy in the clear speech condition and the other three noise conditions did not differ from word recognition accuracy in the baseline speech condition.

There were significant interactions between age and context, $F(1, 390) = 34.21$, $p < .001$, talking condition and

[1] $\theta = \arcsin\sqrt{(X/N)} + \arcsin\sqrt{((X+1)/(N+1))}$
$RAU = (146/\pi)\,\theta - 23$
X denotes the number of correct items and N denotes the total number of items in the test set.

context, $F(6, 390) = 3.73$, $p < .01$, and age and talking condition, $F(6, 390) = 3.54$, $p < .01$. In the high-context condition, older adults performed similarly to younger adults ($p > .05$), but in the low-context condition, the average score for older adults was 14 percentage points lower than the average score for younger adults ($p < .001$). In the high-context condition, word identification scores for speech produced in louder babble were higher (98%) than for baseline speech (94%; $p < .05$), while in the low-context condition, scores for loud speech (81%) were higher than scores for baseline speech (72%; $p < .001$); word recognition scores in other talking conditions were not different from word recognition scores for baseline speech in either context. Relative to the word recognition scores for baseline speech (76%), older adults obtained higher scores when speech was produced loudly (84%) or in any of the noise conditions (82-83%) except for softer speech spectrum noise (p's $< .01$). However, word recognition for clear speech and baseline speech did not differ. In contrast to older adults, younger adults did not correctly recognize more words when speech was produced in noise or loudly, and younger adults correctly recognized fewer words in clear speech (84%) than in baseline speech (90%; $p < .01$).

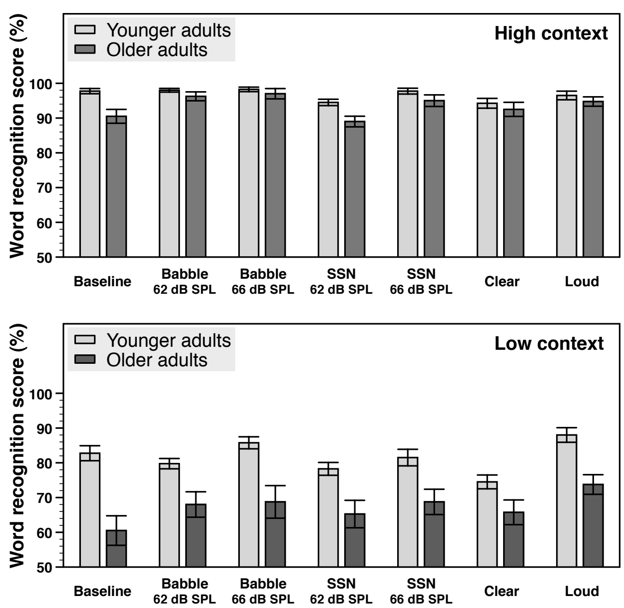

Figure 2. Mean word recognition scores for younger and older adults for high- and low-context sentences, while listening to different talking conditions in a noise environment of 0 dB SNR (SSN = speech spectrum noise). Error bars are standard error of the mean.

2.3 Discussion

Older adults obtained lower word recognition scores overall compared to younger adults, and there was a typical age-related difference when there was a lack of sentence contextual support.

Benefits from loud speech and speech produced in noise were only experienced by older listeners, which may have been due in part to ceiling effects for younger listeners. Surprisingly, clear speech was not more intelligible than baseline speech for either age group, and it was actually less intelligible than baseline speech for younger listeners.

In Experiment 2, we examined whether the benefits of loud speech and speech produced in noise would be more apparent in younger adults in a more difficult listening environment. We tested a different group of younger and older listeners on five conditions (baseline speech, speech produced in babble at 66 dB SPL, clear speech and loud speech) in babble noise with an SNR of -2 dB.

3. EXPERIMENT 2

3.1 Method

3.1.1 Stimuli and procedure

Participants heard five of the eight conditions from Experiment 1: baseline speech, speech in babble at 66 dB SPL, clear speech and loud speech by the new talker, and the original List 1 (which was later excluded from the analyses). Sentences were presented to participants using the same method as in Experiment 1, except that the SNR was -2 dB. The instructions, counterbalancing and scoring procedures were identical to those of Experiments 1, and data analysis was performed similarly as before.

3.1.2 Participants

Participants were 15 younger adults who were undergraduate students (mean age = 21.5 years, SD = 2.4) and 15 community-dwelling older adults (mean age = 68.3 years, SD = 3.5). The criteria for participation were the same as in Experiment 1 (see Figure 1 for participants' average pure-tone audiometric thresholds), and participants had not participated in the previous experiment. Participants gave informed consent and were paid at an hourly rate.

3.2 Results

As shown in Figure 3, there was a significant main effect of age, $F(1, 28) = 5.84$, $p < .05$, with younger adults obtaining higher word recognition scores than older adults. There was a significant main effect of context, $F(1, 196) = 692.49$, $p < .001$, with higher scores for high-context sentences than for low-context sentences. There was also a significant main effect of talking condition, $F(3, 196) = 20.36$, $p < .001$. Multiple t-tests with a Bonferroni correction showed that word recognition scores were not significantly different for speech produced loudly (82%) and speech produced in babble (77%; $p = .07$), while word recognition scores for these conditions were higher than scores for baseline (73%) and clear speech (71%; p's $< .05$). There was no significant difference between word recognition scores for the baseline and clear speech conditions.

There were significant interactions between age and context, $F(1, 196) = 8.53$, $p < .01$; in the high-context condition, younger and older adults performed similarly, but in the low-context condition, the average score for younger adults was 13 percentage points higher than for older adults ($p < .01$). There was a significant interaction between talking condition and context, $F(3, 196) = 7.36$, $p < .001$. For high-context sentences, word recognition scores were higher for speech produced in babble (93%) than for baseline (89%) or clear speech (84%; p's < .05), but the scores were similar for loud speech and speech produced in babble. For low-context sentences, word recognition scores for loud speech (72%) were higher than scores for all other talking conditions by 11 to 15 percentage points (p's < .001), while scores for other conditions were not significantly different. There was a marginal interaction between age and talking condition, $F(1, 196) = 2.62$, $p = .052$. For younger adults, word recognition scores for loud speech (87%) were higher than scores for all other conditions by 7 to 10 percentage points (p's < .01), but scores for other conditions did not differ significantly. For older adults, word recognition scores did not differ significantly for loud speech (77%) and speech produced in noise (74%), but scores for both conditions were higher than for baseline (69%) and clear speech (66%; p's < .05). Word recognition did not differ for baseline and clear speech.

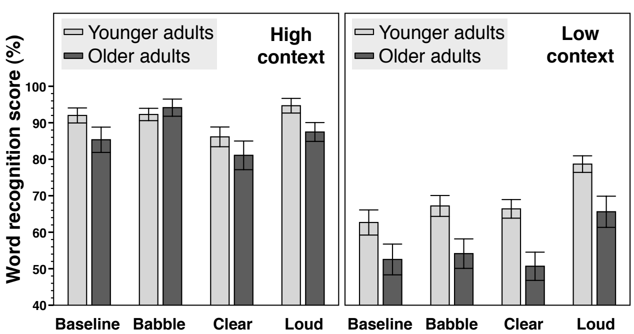

Figure 3. Mean word recognition scores for younger and older adults for high- and low-context sentences, while listening to different talking conditions in a noise environment of -2 dB SNR. (Base = baseline condition). Error bars are standard error of the mean.

3.3 Discussion

Similar to Experiment 1, younger adults obtained higher word recognition scores overall compared to older adults. In contrast to Experiment 1, younger adults benefited from loud speech, but clear speech was not different from the baseline condition. Older adults found both loud speech and speech produced in noise helpful.

Since listening difficulty was not equated between the two age groups in Experiments 1 and 2, the differences between age groups in their response to different talking conditions might have been due to a simple effect of listening difficulty, rather than differences in how younger and older adults used acoustic cues in different talking conditions. Therefore, in Experiment 3, we tested both age groups at their SPIN thresholds (i.e., the SNR at which participants would obtain a 50% correct score on low-context sentences) using the same subset of conditions from Experiment 2.

4. EXPERIMENT 3

4.1 Method

4.1.1 Stimuli and procedure

Participants were seated in a double-walled sound-attenuating IAC booth. To obtain the SPIN threshold of each participant, the original SPIN-R List 5 was presented at a high SNR (usually +6 dB for older adults and +4 dB for younger adults) and the original SPIN-R List 3 was presented at a low SNR (usually 0 dB for older adults and -1 dB for younger adults). List 6 was used for additional testing if the participant performed unexpectedly well or poorly in the first SNR condition, and linear interpolation was used to calculate the SPIN threshold. Younger adults had an average SPIN threshold of +2 dB SNR, whereas older adults had an average SPIN threshold of +4 dB SNR.

After each participant's SPIN threshold was obtained, the participant was tested on the same five lists as in Experiment 2 (with List 1 excluded from analyses). Sentences were presented to participants using the same method as in Experiments 1 and 2, except that the SNR was the participant's SPIN threshold. The instructions and counterbalance and scoring procedures were identical to those of Experiments 1 and 2, and data analysis was performed similarly as in Experiments 1 and 2.

4.1.2 Participants

Participants were 15 younger adults who were undergraduate students (mean age = 20.8 years, SD = 2.1) and 15 community-dwelling older adults (mean age = 71.3 years, SD = 5.8). The criteria for participation were the same as in Experiment 1 and 2 (see Figure 1 for participants' average pure-tone audiometric thresholds), and participants had not participated in either of the previous experiments. Participants gave informed consent and were paid at an hourly rate.

4.2 Results

As shown in Figure 4, there was no significant main effect of age, $F(1, 28) = 2.55$, $p = .1$. There was a significant main effect of context, $F(1, 196) = 558.77$, $p < .001$, with higher word recognition scores for high-context sentences than for low-context sentences. There was also a significant main effect of talking condition, $F(3, 196) = 6.68$, $p < .001$. Multiple *t*-tests with a Bonferroni correction showed that word recognition scores were higher for loud speech (92%) than for baseline speech (88%) and speech produced in babble (89%; p's < .01) and marginally more intelligible than clear speech (89%; $p = .07$), but word recognition scores for other talking conditions did not differ significantly. There was a significant interaction between age and context, $F(1, 196) = 26.16$, $p < .001$; in the high-context condition, older adults performed similarly to younger adults, but in the low-context condition, the average word recognition score for older adults was lower than that of younger adults by 8 percentage points ($p < .001$). There was a significant interaction between talking condition and context, $F(3, 196) = 4.21$, $p < .01$. For high-context sentences, there were no significant differences between talking conditions, but for low-context sentences, word recognition scores were higher for loud speech than for all other conditions by 6 to 8 percentage points (p's $< .01$), though word recognition did not differ in other talking conditions. There was no significant interaction between age and talking condition, $F(3, 196) = 1.29$.

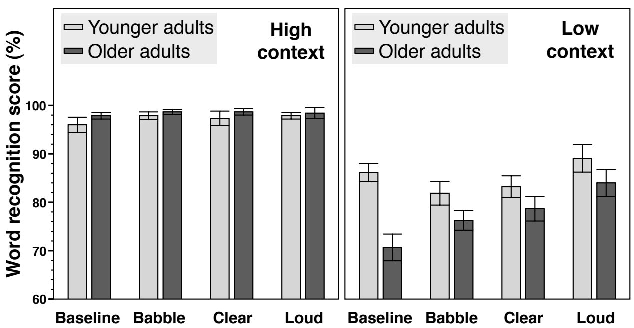

Figure 4. Mean word recognition scores for younger and older adults for high- and low-context sentences, while listening to different talking conditions at their SPIN thresholds. (Base = baseline condition). Error bars are standard error of the mean.

4.3 Discussion

Unlike in Experiments 1 and 2, there was no overall age-related difference in word recognition scores after listening difficulty was equated for younger and older adults, although younger adults still outperformed older adults in the low-context condition. Both age groups benefited from loud speech but only in the low-context condition.

5. ACOUSTIC ANALYSIS

5.1 Procedure

Speech stimuli were measured using the PRAAT speech analysis program (Boersma & Weenink, 2012). The annotation function in PRAAT was used to mark the boundaries of the sentence, the final keyword and the location of the vowel or dipthong within the word. Measures of duration, F_0 and intensity were extracted at each level using a custom script. Variation in intensity (intensity SD) and variation in speaking F_0 (F_0SD) were measured for sentences, and F1 and F2 were measured for vowels. As the sentence lists were similar in their distributions of different vowels and dipthongs, values for each formant were averaged within each talking condition. A subset (14%) of the stimulus files was independently annotated by two researchers, and the differences between measures obtained by the second researcher were within ±7% of the absolute values of measures obtained by the first.

5.2 Data analysis

We focused our analyses on four main talking conditions that led to differences in word recognition: baseline speech, speech produced in babble at 66 dB SPL, clear speech and loud speech. Table 2 shows the acoustic measures taken for sentences and sentence-final target words. For both sentences and words, individual acoustic measures were modelled as a function of talking condition (four conditions), context and the interaction between these two factors, using a random intercept model with a compound symmetry covariance structure.

5.3 Results

For sentences, there was a significant main effect of talking condition on speaking rate, $F(3, 192) = 17.42$, $p < .001$, with clear speech having a slower speaking rate than baseline speech and loud speech (p's $< .001$). There was also a significant main effect of talking condition on mean F_0, $F(3, 192) = 271.56$, $p < .001$; clear speech had a lower F_0 and loud speech had a higher F_0 than any other talking condition (p's $< .001$). However, F_0SD did not differ significantly between conditions. There was a significant main effect of talking condition on mean intensity, $F(3, 192) = 16.59$, $p < .001$, an interaction of talking condition with context for intensity, $F(3, 192) = 3.44$, $p < .05$, and an interaction of talking condition with context for intensity SD, $F(3, 192) = 2.78$, $p < .05$. However, the largest overall intensity difference between conditions was 0.3 dB, and the largest difference in intensity SD was 0.7 dB; therefore, overall sentence-level differences in intensity are unlikely to be of practical significance. There was no main effect of context on any acoustic measure.

For target words, there was no significant effect of talking condition on word duration or formant values, but there was a significant effect of talking condition on F_0, $F(3, 192) = 24.37$, $p < .001$, such that words spoken loudly had a higher F_0 than words spoken clearly or in babble (p's $< .01$), while words spoken clearly had a lower F_0 than in all other conditions (p's $< .001$). There was also a significant main effect of intensity, $F(3, 192) = 11.36$, $p < .001$; words spoken loudly had a higher intensity than in all other conditions (p's $< .01$), but words did not differ on intensity among other talking conditions. There was no significant effect of context on any acoustic measure, and no interaction of talking condition with context.

Table 2. Mean acoustic measures of sentences and target words from four conditions, with standard deviations in parentheses (Base = baseline condition; Babble = babble noise at 66 dB SPL).

	Base	Babble	Clear	Loud
Sentences				
Rate (syl/s)	4.1 (0.4)	3.9 (0.4)	3.7 (0.4)	4.3 (0.5)
F_0 mean (Hz)	124 (9)	123 (7)	113 (5)	153 (7)
F_0SD (Hz)	25.7 (19.5)	27.2 (15.2)	21.5 (13.5)	25.1 (6.8)
Intensity (dB)	68.8 (0.3)	69.0 (0.2)	68.9 (0.2)	69.1 (0.2)
Intensity SD (dB)	10.9 (1.2)	11.1 (1.3)	10.9 (0.9)	11.0 (1.2)
Target words				
Duration (s)	519 (94)	547 (129)	557 (88)	521 (94)
F_0 mean (Hz)	122 (25)	117 (27)	99 (12)	132 (12)
Intensity (dB)	66.8 (1.9)	66.9 (1.8)	66.2 (1.7)	68.2 (1.6)
F1 (Hz)	619 (124)	580 (132)	580 (139)	617 (135)
F2 (Hz)	1442 (358)	1522 (402)	1540 (391)	1544 (364)

5.4 Discussion

Although sentences were equated on overall intensity and intensity variability was similar between talking conditions, the distribution of energy was different between talking conditions. In loud speech, the intensity of sentence-final target words was about 2 dB higher than in other conditions, which likely contributed to the higher word recognition scores in the loud speech condition. There were some benefits of speech produced in noise, but there were no clear acoustic differences between speech produced in noise and baseline speech. Listeners did not benefit from clear speech, which was characterized by a lower mean F_0 and a slower speaking rate compared to other talking conditions. In general, talkers produce clear speech using a slower speaking rate than in conversational speech; however, studies have suggested that speaking rate may not be the most important factor that affects intelligibility (Krause & Braida, 2002; Krause & Braida, 2004). Other properties such as F_0 and formant frequency may interact with the noise environment and listener characteristics to affect word recognition. One study found that F2 was raised in clear speech, which led to poorer vowel recognition by listeners with high-frequency hearing loss, presumably because the acoustic information became less audible (Ferguson & Kewley-Port, 2002). In our study, the mean F_0 of target words in clear speech was about 20 Hz lower than in other conditions. Since F_0 is an important cue for segregating a target from background noise (Oxenham, 2008), the lower F_0 of the target words may have enabled the background noise to mask them more effectively, resulting in lower word recognition scores.

6. SUMMARY OF FINDINGS

6.1 Effect of listener age

Younger and older adults listened to sentences that had been recorded in different talking environments, which were then played in different levels of background noise. Younger adults correctly identified more words than older adults in noise environments in which listening difficulty was not equated for the two age groups, but the two groups did not perform differently when tested at their respective SPIN thresholds. In all noise environments, the two age groups performed similarly for high-context sentences, but younger adults obtained higher scores than older adults for low-context sentences.

6.2 Effect of speaking condition

Loud speech was always more intelligible than baseline speech; however, in the two most difficult listening environments, speech produced in louder babble was also more intelligible than baseline speech. Clear speech did not lead to better word recognition in any of the tested conditions. For high-context sentences, there were minimal differences between talking conditions, but for low-context sentences, loud speech was always beneficial. When listening difficulty was not equated for younger and older adults, the two age groups differed on which talking conditions they found more intelligible in noise. Younger adults found loud speech more intelligible than baseline speech only in the most difficult listening environment, whereas speech produced in babble at 66 dB SPL was no different than baseline speech. In contrast, older adults found both speech produced in babble at 66 dB SPL and loud speech more intelligible than baseline speech. When tested at their respective SPIN-R thresholds, both younger and older adults found loud speech, but not speech produced in babble, more intelligible than baseline speech. The higher intensity of target words in loud speech and the lower F_0 of target words in clear speech may explain the benefit of loud speech over clear speech.

7. CONCLUSION

The purpose of the study was to investigate whether word recognition performance would improve on a standard test of word recognition when the condition in which the speech stimuli were produced were matched to the listening environment, and to investigate whether younger and older adults would benefit differently. Loud speech improved word recognition in noise but clear speech was not helpful; for listeners with normal or near-normal hearing, intensity changes in speech may be more important than changes in articulation when listening to speech in a noisy environment. The benefits of speech produced in noise were less consistent compared to the benefits of loud speech; it is possible that these benefits would have been more evident if the talker had been exposed to higher noise levels than those used in this study. Older adults with good hearing benefited more than younger adults when listening to speech that was matched to the noise environment, but only under conditions in which listening difficulty was greater. Therefore, listener performance may change on a test of speech intelligibility when speech is produced under more ecologically valid circumstances. However, these effects may be more apparent in older adults than in younger adults and when task difficulty increases due to a lack of supportive contextual information or an increase in background noise.

ACKNOWLEDGMENTS

This study was funded by a Natural Sciences and Engineering Council of Canada grant awarded to M. Kathleen Pichora-Fuller. The authors wish to thank James Qi for technical assistance, and Sabrena Deonarain and Elizabeth Olsewska for their assistance in recruiting participants.

REFERENCES

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, *27*(1), 32-48.

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.35, retrieved 8 December 2012 from http://www.praat.org/

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*, *76*(1), 87-96. doi:10.1121/1.391011

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, *112*(1), 259-271. doi:10.1121/1.1482078

Forrest, K., Abbas, P. J., & Zimmermann, G. N. (1986). Effects of white noise masking and low pass filtering on speech kinematics. *Journal of Speech and Hearing Research*, *29*(4), 549-562.

Huang, Y., Xu, L., Wu, X., & Li, L. (2010). The effect of voice cuing on releasing speech from informational masking disappears in older adults. *Ear and Hearing*, *31*(4), 579-583. doi:10.1097/AUD.0b013e3181db6dc2

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61(5), 1337-1351. doi:10.1121/1.381436

Krause, J.C., and Braida, L.D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. Journal of the Acoustical Society of America, 112 (5), 2165-2172. doi:10.1121/1.1509432

Krause, J.C., and Braida, L.D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. Journal of the Acoustical Society of America, 115 (1), 362-378. doi:10.1121/1.1635842

Letowski, T., Frank, T., & Caravella, J. (1993). Acoustical properties of speech produced in noise presented through supra-aural earphones. *Ear and Hearing*, 14(5), 332-338. doi:10.1097/00003446-199310000-00004

Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: Implications for hearing loss and cochlear implants. *Trends in Amplification*, *12*(4), 316-331. doi:10.1177/1084713808325881

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, *28*(1), 96-103.

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, *47*(SUPPL. 2), S72-S82. doi:10.1080/14992020802307404

Pittman, A. L., & Wiley, T. L. (2001). Recognition of speech produced in noise. *Journal of Speech, Language, and Hearing Research*, *44*(3), 487-496. doi:10.1044/1092-4388(2001/038)

Schum, D. J. (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, *7*(3), 212-218.

Sherbecoe, R. L., & Studebaker, G. A. (2004). Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. *International Journal of Audiology*, *43*(8), 442-448. doi:10.1080/14992020400050056

Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Linguistics and Language Compass*, *3*(1), 236-264. doi:10.1111/j.1749-818X.2008.00112.x

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America*, *84*(3), 917-928. doi:10.1121/1.396660