Across Talker Variability in Auditory, Visual, and Audiovisual Speech Intelligibility for Conversational and Clear Speech

Jean-Pierre Gagné
École d'orthophonie et d'audiologie
Université de Montréal

Valerie Masterson
Hearing Health Care Research Unit
Department of Communicative Disorders
The University of Western Ontario

Kevin G. Munhall
Department of Psychology
Queens University

Nancy Bilida
Department of Audiology
St-Josephs Health Center
Hamilton, Ontario, Canada

Carol Querengesser
Hearing Health Care Research Unit
Department of Communicative Disorders
The University of Western Ontario

Ten female adults were recorded while saying a list of test words twice: once using conversational speech and once using clear speech. The recorded stimuli were randomized and presented to groups of subjects in one of three sensory modalities: degraded auditory, visual-alone, or degraded audiovisual. For the auditory and the audiovisual conditions, a hearing loss simulator (HELOS: Gagné & Erber, 1987) was used to simulate the effects of a sensorineural hearing loss. The perceptual data were used to determine the speech intelligibility of individual talkers under each experimental condition. The results revealed that, in each sensory modality, there was a significant difference in speech intelligibility among the talkers. A comparison of speech intelligibility scores obtained for the tokens of conversational and clear speech revealed that there was a significant interaction between talkers and manner of speech. Many talkers did not produce a significant clear speech effect in any sensory modality. Others produced a clear speech effect in one but not all sensory modalities.

Communication partners have an important effect on the speech perception performance of individuals with hearing impairment. Anecdotal reports obtained from individuals with hearing impairment suggest that some talkers are easier to *understand* than others. Also, there exists a body of experimental literature that indicates variability in speech intelligibility across talkers. Speech intelligibility refers to aspects of speech-language output that allow a communication partner to understand what the talker is saying (Carney, 1986). Studies have been conducted on both *auditory* and *visual* speech intelligibility. For example, in a recent study by Cox, Alexander, and Gilmore (1987), words embedded in a carrier phrase (subtests of the Speech Perception of Auditory Contrast test – SPAC test – Boothroyd, 1985) were used to investigate differences in *auditory* speech intelligibility among 3 male and 3 female talkers. The speech stimuli were recorded in various everyday acoustic environments. The recorded materials were presented to groups of normal-hearing subjects. The authors reported significant differences in speech intelligibility across talkers. For recordings obtained in a relatively quiet listening environment (i.e., +7 dB speech-to-multitalker babble ratio), the range of speech intelligibility scores across talkers was 14%. When additional noise was added to the same recordings, in order to eliminate ceiling effects in the speech recognition scores, the range of speech intelligibility scores across talkers was approximately 24%. Other investigators have also reported significant differences in auditory speech intelligibility across talkers (e.g., Creelman, 1957; Hood & Poole, 1980; Kreul, Bell, & Nixon, 1969; Mullennix, Pisoni, & Martin, 1989; Palmer, 1955; Penrod, 1979; Verbrugge, Strange, Shankweiler, & Edman, 1976; Weinhouse & Miller, 1963; Williams & Hecker, 1968). Differences in *visual* speech intelligibility have also been reported for the production of consonants, vowels, and sentences (e.g., Grant & Braida, 1991; Kricos & Lesner, 1982; Lesner & Kricos, 1981). For example, in one study 6 talkers (3 females judged to be difficult to speechread and 3 females judged to be fairly easy to speechread) were videotaped while saying 24 consonants (/a-*C*-a/) and a set of sentences (Kricos & Lesner, 1982). For each type of stimuli, a test tape consisting of a randomization of the recordings obtained from each talker was prepared and shown (visually-only) to a group of normal-hearing subjects. The visual consonant recognition scores obtained ranged from 61% (for the least intelligible talker) to 78% (for the most intelligible talker).

Clear speech has been defined as that speech which naturally arises when a talker is in a difficult communication situation (Picheny, Durlach, & Braida, 1985). The acoustic characteristics and the concomitant perceptual effects of clear speech on *auditory* speech perception have been investigated (e.g., Chen, 1980; Moon & Lindblom, 1989; Payton, Uchanski, & Braida, 1994; Picheny et al., 1985; Picheny, Durlach, & Braida, 1986, 1989; Uchanski, 1988). The results of those investigations revealed that some acoustic differences between the properties of conversational and clear speech exist at the phonetic, phonological, and sentence levels. Perceptually, the use of clear speech can result in substantial increases in the auditory speech recognition performances of listeners. For example, Picheny et al. (1985) reported that, for some talkers, the use of clear speech resulted in improvements of more than 17% among a group of 5 subjects with sensorineural hearing impairment.

These findings have promising applications for audiologic rehabilitation. For example, the speech perception difficulties of individuals with a hearing impairment could be reduced by instructing frequent communication partners to use clear speech (e.g., De Filippo, 1991; Erber, 1988, 1993; Kaplan, Bally, & Garretson, 1985; Tye-Murray & Schum, 1994). However, several issues related to the effects of clear speech on speech perception remain to be investigated. Among them are the effects of clear speech on visual and audiovisual speech perception. This question is important because in many communication settings, individuals with hearing impairment rely on the visual as well as the audio signal to perceive speech.

There are few data pertaining to the effects of clear speech on visual and audiovisual speech perception. Franks (1979) investigated the effects of exaggerated mouth movements on speechreading performance. A female talker produced a series of consonants, spondees, and sentences (live-voice presentations) using both conversational and exaggerated speech. The stimuli were shown to a group of subjects with normal hearing and individuals with a sensorineural hearing loss. The results revealed that exaggerated speech had (a) a deleterious effect on visual consonant recognition scores, (b) no effect on visual spondee recognition scores, and (c) a significant enhancement effect on the visual sentence recognition scores. Stoker and French-St. George (1987) investigated the effects of training on the lipreadability of talkers. CID everyday sentences were recorded from 13 talkers who completed a 36-hr course designed to train oral interpreters. The course included lessons on how to produce visually intelligible speech. The talkers were videotaped before and after the completion of the course. During the recordings, the talkers were instructed to articulate as if they were talking to someone who had a hearing impairment. The speech samples were randomized (across talkers and recording sessions) and shown to normal-hearing subjects and persons with a hearing impairment. The results of the perceptual experiment revealed that there was substantial variability in visual speech intelligibility across talkers (i.e., the range of scores exceeded 30%) and the rank-ordering of the talkers' speech intelligibility changed between the pre- and post-training recordings. Based on the perceptual data, the majority of talkers were more intelligible (visually) after they completed the oral interpretation training program. The subjects displayed an average improvement of approximately 8% in visual sentence recognition for the sentences recorded during the post-training session. However, the amount of improvement in speech intelligibility displayed by individual talkers varied considerably. Some talkers displayed no improvement in visual speech intelligibility pre- and post-training and others showed a substantial amount of improvement in speech intelligibility (as much as 17%) post-training. In general, the talkers who were the most intelligible visually based on the pre-training recordings displayed less improvement in visual speech intelligibility post-training.

The long-term objective of our research program is to identify the articulatory (kinematic) characteristics of clear speech so that they can be taught to persons who display less than optimal speech intelligibility when they communicate with individuals who have a hearing impairment. The present study was designed: (a) to examine across talker variability in auditory, visual, and audiovisual speech intelligibility for words, and (b) to describe the effects of clear speech on the auditory, visual, and audiovisual perception of words.

METHOD

Talkers

The talkers were 10 caucasian female adults. They were faculty members ($n = 1$), secretarial staff ($n = 7$), and students ($n = 2$) in the Faculty of Applied Health Sciences at the University of Western Ontario. All the talkers were recruited on the basis of their availability and their willingness to take part in the investigation. In addition, all the talkers met the following selection criteria: (a) English as their first language; (b) no atypical speech characteristics nor any noticeable oral-facial abnormalities, as judged by two of the experimenters (J-PG and VM); and (c) no pronounced regional accent, as judged by the same two experimenters. Prior to the investigation no data were available on the speech intelligibility of the talkers. The talkers were not remunerated for their participation and they were not given any details concerning the purpose of the study.

Test Stimuli

Selection of stimulus set. The present study was designed to investigate speech intelligibility for tokens of conversational and clear speech in different sensory modalities. However, a long-term goal of the present research program is to investigate the differences in articulatory patterns (i.e., the kinematics) that characterize conversational and clear speech. The selection of a stimulus set that would make it possible to highlight differences in the *kinematics* of conversational and clear speech was a major challenge. A decision was made to select, as much as possible, speech stimuli that approximated speech tokens used in everyday conversations. However, it is well known that speech recognition scores are influenced greatly by the linguistic redundancies and contextual cues available in the stimulus set (e.g., Boothroyd, 1988; Gagné, Tugby, & Michaud, 1991; Kalikow, Stevens, & Elliott, 1977). We sought speech stimuli that would minimize the effects of linguistic and contextual cues in order to focus on the articulatory aspects of speech. One alternative was to use sentences that were devoid of contextual cues (e.g., Picheny et al., 1985). However, this option was rejected because one of the experimental conditions under investigation consisted of a visual speech recognition task. It was concluded that nonsense sentences would be too difficult to speechread and would potentially lead to floor effects in the visual speech recognition scores. Previous investigations have shown that durational and other temporal cues constitute a major difference between conversational and clear speech. Among other features, clear speech is characterized by longer duration of utterances as well as more pauses and longer pauses between words (Picheny et al., 1986, 1989; Uchanski, 1988). Given our future plans to focus on the kinematic characteristics of speech production, we sought speech stimuli that would minimize the contribution of temporal cues between tokens of conversational and clear speech. The effects of durational (and other temporal cues) are more evident in sentence length stimuli than in words presented in isolation (Uchanski, 1988). Hence, the stimulus set selected for the present investigation consisted of mono- and bi-syllabic words spoken in isolation.

In order to avoid the possibility of floor effects in the visual-only perceptual task, a pilot investigation was conducted to identify test words that would yield mean visual word recognition scores of approximately 60% correct for tokens of conversational speech. Two talkers, who did not participate in the actual study, were videotaped while saying a total of 25 monosyllable and 30 spondaic words chosen from the CID-W1, W2, and W-22 word lists (Davis & Silverman, 1970). The same recording procedures as those employed in the actual study (described below) were used in the pilot investigation. The recorded stimuli were randomized and shown to a group of normal-hearing subjects who did not take part in the actual investigation. An item analysis was performed and test stimuli that yielded the desired overall level of performance were retained. The 14 test stimuli selected for the study (8 monosyllables and 6 bisyllables) appear in Appendix A.

Test stimuli. The talkers were videotaped in a 9 ft × 9 ft double-walled audiometric test suite. A Sony 8 mm color video camera recorder (CCD V220) was placed approximately 1.5 m directly in front of the talker. In order to focus on the effects of articulatory patterns on visual speech intelligibility, only the lower half of the talker's face (extending from the nose to just below the chin) was recorded, thus minimizing the effects of facial expression on speech intelligibility. None of the talkers wore lipstick or facial make-up. Lights were used to reduce shadows on the talker's face; however, no special attempt was made to illuminate the inside of the talker's mouth. A lapel-microphone connected directly to the video camera was used to record the talker's utterances. A calibration tone (a 1-kHz pure tone whereby 0-dB VU meter = 84 dB SPL) was recorded at the beginning of the videotape.

During the recording sessions, the talkers sat directly in front of the video camera. The lapel microphone was attached to their clothing, at a distance of 30 cm from their mouth. For each talker, the recording equipment was adjusted and instructions were given to the talker. Each talker completed a few practice items before the actual recording session began. During the recording session, the talker was instructed to wait for a visual cue from the experimenter before she spoke a test word while looking into the lens of the camera. In addition, she was asked to start and finish each utterance from a closed-mouth position. In some cases a test word had to be repeated, either because the talker was not positioned properly (i.e., the visual image was not centered on the viewing monitor), she did not start or finish the utterance from a closed-mouth position, or she did not follow other instructions given by the experimenter.

During the initial recording of the stimulus set, the talkers were not given any instructions concerning the manner in which they were to produce the utterances. Presumably they were using articulation patterns that approximated conversational speech. Following the initial recording of the complete stimulus set, each talker was asked to say the same words a second time, using clear speech. She was not given specific instructions concerning the characteristics of clear speech. The talker was instructed to "articulate each word clearly as if you were communicating with someone who had difficulty understanding what you are saying." Throughout the second recording session the talkers were reminded to articulate clearly. The recordings obtained from each of the talkers were used to prepare the test tape.

An 8 mm editing unit (Sony EVO-720) was used to edit the recordings from the master tape. The test items were completely randomized (across talkers and manner of speech). The test tape consisted of 280 test items (14 stimuli × 10 talkers × 2 manners of speech). Each test item consisted of a 3 s test-item identification number, the stimulus word, and a 6 s written message that appeared on the TV monitor and prompted the subjects to provide a response.

Subjects

A total of 48 subjects took part in the investigation. More than 85% of the subjects were students in the Department of Communicative Disorders at The

University of Western Ontario (London, Ontario, Canada). The remaining subjects were either friends of students from that Department or students in other academic programs at the same University. Subjects ranged from 18-35 years of age, had normal hearing (hearing detection thresholds better than 25 dB HL [re: ANSI-1969] at .5, 1, 2, and 4 kHz) and normal (or corrected normal) visual acuity (20/20 binocular vision as measured with a Snellen chart). Initially 24 subjects were recruited for the visual-alone word recognition task. Later, another group of 24 subjects were recruited for the investigation. Of this latter group, 12 subjects were randomly assigned to the degraded auditory word recognition task and the other 12 subjects were assigned to the degraded audiovisual word recognition task. None of the subjects completed the task in more than one sensory modality. The subjects volunteered for the experiment and they were not remunerated for their participation in the study.

Procedure

The test procedure varied slightly according to the experimental condition under investigation. In general, the subjects were tested in small groups (not exceeding 4 individuals per session). To reduce learning and familiarization effects, the starting position of the test tape was staggered. In each sensory modality, an equal number of subjects started the perceptual task at test item 1, 71, 141, or 211. Before the initial test session, the subjects were given 5 practice items (those items were not included in the test proper). They were instructed to attend to each test stimulus and to indicate their response at the appropriate place on a prepared response form. The subjects were instructed to supply an answer for each test item and to guess if they were not certain of the correct response. The subjects were tested in one or two sessions for a total of approximately 2 hr. Whenever the testing was completed in one session the subjects were given a 15-min rest period midway through the test session.

Visual word recognition task. Pilot investigations revealed that, for the visual word recognition task, the overall performance of the subjects was less than 25% correct when the test was completed in an open-set response format. To eliminate the possibility of floor effects, the task was changed to a four-alternative closed-set response task for this experimental condition only. The response foils selected were similar to the test words (see Appendix A for a list of the test stimuli and foils). However, the three foils differed from the test words in at least one consonant viseme (based on the universal consonant viseme categories reported by Jackson, 1988). A series of preliminary visual word recognition experiments were conducted to identify response foils for each of the test words that would yield overall visual word recognition scores of approximately 60-70% correct. For this experimental condition, the subjects were instructed to circle their response at the appropriate place on a response form.

The stimuli were displayed on a 33 cm color TV monitor. The subjects were seated approximately 1 m from the monitor, at an angle that permitted an undistorted view of the video image (i.e., less than 45° from the center of the screen). For this task the audio signal was not presented to the subjects.

Degraded auditory and degraded audiovisual word recognition tasks. For the auditory-alone and the audiovisual experimental conditions, a hearing loss simulator (HELOS) was used to degrade the audio signal in order to simulate aspects of a sensorineural hearing loss and to eliminate ceiling effects in the word recognition scores. The signal processing strategies incorporated into HELOS and the perceptual effects observed under different hearing loss simulation conditions were consistent with those described by Gagné and Erber (1987). Pilot investigations were conducted to determine the threshold setting of HELOS and the level of the test stimuli required to ensure: (a) that the subject's overall level of performance would not exceed approximately 80% correct recognition scores in the audiovisual mode and, (b) that all the words were audible without being presented at levels that would be uncomfortably loud for the subjects. For the HELOS device used in the present investigation the desired perceptual effects were achieved for the following HELOS settings: (a) flat audiometric configuration, (b) distortion level knob set to zero, and (c) detection threshold knob set to position 10 (Bilida, 1992). Thus, in the present investigation there was no attempt to simulate sloping audiometric configurations or to simulate the effects of reduced frequency and temporal resolution. Only center-clipping was used to degrade the audio signal. According to Gagné and Erber (1987), center-clipping simulates two aspects of sensorineural hearing loss: loss of hearing sensitivity and loudness recruitment. Once the hearing loss simulation settings were determined, the equipment was calibrated and the same settings were used for both the degraded auditory and the degraded audiovisual word recognition tasks.

The perceptual tests were conducted in the same audiometric test suite that was used for recording the talkers. The audio signal was delivered via a small loudspeaker (Paradigm Titan) that was placed on top of a 66 cm color TV monitor (Samsung, model CT-680). A free-field calibration of the audio signal was completed before each testing session. The subjects sat approximately 2 m from the TV monitor at an angle that permitted an undistorted view of the visual image (i.e., less than 45° from the center of the screen). For the degraded auditory experimental condition, the video output of the videotape player (Sony 8 mm EV-C3) was disconnected.

An open-set response format was used for the degraded auditory and the degraded audiovisual word recognition tests. The subjects were informed that the stimuli consisted of words presented in isolation. They were not told that the stimuli consisted of repeated presentations of the same 14 test words. Each subject was instructed to write the test word at the appropriate place on the response form following each stimulus presentation.

Scoring

For the visual-alone perceptual task the subjects were credited with a correct response only if they circled the correct test word on the response form. For the degraded auditory and the degraded audiovisual conditions the subjects were given credit for a correct response only if they wrote the exact test word at the appropriate place on the response form. For the latter two experimental conditions, inserted or deleted morphemes (such as *s*) were considered to be incorrect responses.

RESULTS

The perceptual data were used to investigate the conversational speech intelligibility of individual talkers and to examine the effects of clear speech in each sensory modality. In the present investigation, a talker's speech intelligibility was determined by computing the percent correct word recognition score obtained by the groups of subjects for a given experimental condition. Because different scoring methods and response formats were used for the degraded auditory, visual-alone, and degraded audiovisual perceptual tasks, the results obtained in each sensory modality were analyzed separately. Also, statistical comparisons of the speech intelligibility scores obtained in the different sensory modalities were not performed because of the differences among the experimental conditions and the fact that, in each sensory modality, the data were obtained from a different group of subjects.

Conversational Speech Intelligibility

To determine the conversational speech intelligibility of individual talkers (in each sensory modality), the data obtained for the two types of stimuli (mono- and bi-syllables) were combined. Also, only the tokens of conversational speech were considered in the analysis. The speech intelligibility scores obtained for each talker are displayed in Figure 1. The degraded auditory speech intelligibility scores (displayed in the top panel) ranged from 10% (talker J) to 40% (talker A). The visual speech intelligibility scores (displayed in the middle panel) ranged from 61% (talker E) to 84% (talker G). The degraded audiovisual speech intelligibility scores (displayed in the bottom panel) ranged from 56% (talker J) to 77% (talker C). Three separate one-way ANOVA revealed that there were significant differences in conversational speech intelligibility among the talkers in each of the sensory modalities under investigation (see section under intelligibility in Table 1).

A visual examination of the data suggests that there does not appear to be a simple relationship in speech intelligibility across sensory modalities. Some talkers maintained their relative ranking in speech intelligibility across sensory modalities. For example, talker J displayed relatively poor speech intelligibility in each sensory modality and talker C displayed higher than average speech intelligibility scores in each sensory modality. However, some talkers who exhibited relatively high speech intelligibility in one sensory modality were ranked poorly, relative to other talkers, in another sensory modality. For example, talker B was

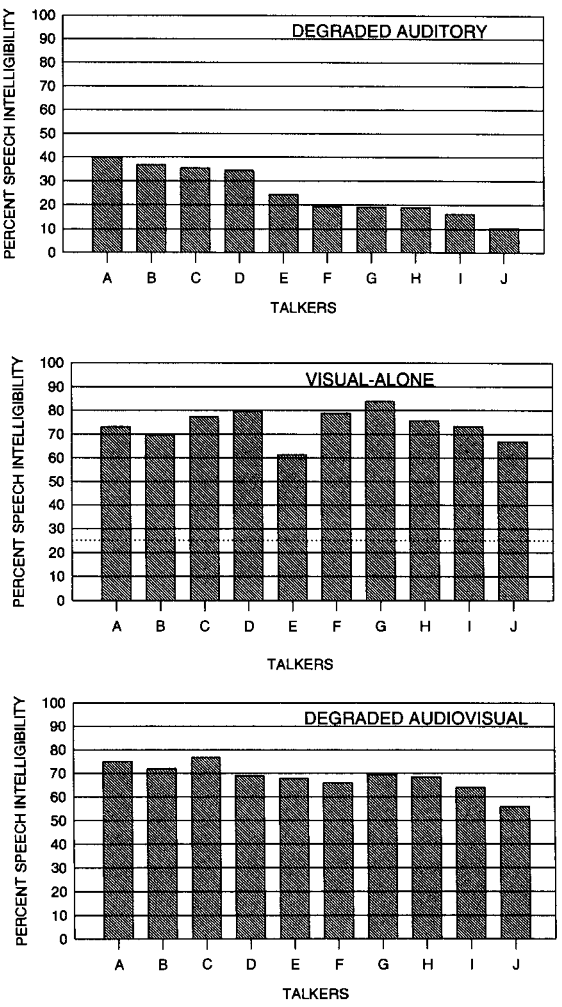

Figure 1. Percent speech intelligibility scores for tokens of conversational speech (monosyllables and bisyllables combined). Data are shown for 10 female talkers, in three different sensory modalities: degraded auditory (top panel), visual-alone (middle panel), and degraded audiovisual (bottom panel). In each sensory modality, the speech intelligibility scores are based on data obtained from a different group of subjects.

Table 1

Results of the Analyses of Variance With Repeated Measures for the Speech Intelligibility Data Obtained for a Group of 10 Female Talkers

Source	*df*	Sum of squares	Mean square	*F* ratio	Significance Prob > *F*
SPEECH INTELLIGIBILITY					
Degraded audio	9	11383.75	1264.87	20.02	.001
Visual-alone	9	9499.42	1055.49	11.3101	.001
Degraded audiovisual	9	3727.83	414.20	4.96	.001
CLEAR SPEECH					
WORD TYPE					
Degraded audio	11	84217.95	84217.9	100.31	.001
Visual-alone	23	.35	.35	.02	.889
Degraded audiovisual	11	4046.65	4046.65	38.22	.001
MONOSYLLABLES					
Degraded audio					
Manner (M)	11	940.10	940.10	9.34	.01
Talker (T)	9	3513.02	390.34	4.58	.001
M × T	9	2705.73	300.64	3.84	.001
Visual-alone					
Manner (M)	23	130.2i	130.21	1.51	.231
Talker (T)	9	8697.92	966.44	6.58	.001
M × T	9	2968.75	329.86	2.45	.01
Degraded audiovisual					
Manner (M)	11	2041.67	2047.67	17.89	.001
Talker (T)	9	4416.67	490.74	2.92	.005
M × T	9	7151.04	794.56	5.18	.001
BISYLLABLES					
Degraded audio					
Manner (M)	11	7593.41	7593.41	16.60	.005
Talker (T)	9	30961.08	3440.12	15.76	.001
M × T	9	8598.58	955.40	5.15	.001
Visual-alone					
Manner (M)	23	3703.52	3703.52	12.09	.005
Talker (T)	9	24479.73	2719.97	12.23	.001
M × T	9	4757.06	528.56	2.83	.005
Degraded audiovisual					
Manner (M)	11	374.95	374.95	2.23	.163
Talker (T)	9	5781.63	642.40	3.64	.001
M × T	9	4116.13	457.35	2.75	.01

ranked second in degraded auditory speech intelligibility but ranked eighth in visual speech intelligibility. On the other hand, talker G was the most intelligible talker visually but she was ranked seventh in degraded auditory speech intelligibility. To a certain extent a talker's degraded audiovisual speech intelligibility tended to reflect her combined ranking in degraded auditory and visual speech intelligibility. For example, talker C was above average in all three sensory modalities. Talker G displayed poorer than average degraded auditory speech intelligibility, better than average visual speech intelligibility, and slightly better than average degraded audiovisual speech intelligibility.

Clear Speech

A main purpose of the present study was to investigate a talker's speech intelligibility for conversational and clear speech, in each sensory modality. Hence, primary variables of interest were: *talkers* and *manner of speech* (as well as the interactions between those two variables). Preliminary analyses revealed that although there was not a significant main effect for word-type (monosyllables vs. bisyllables) when the syllables were presented in the visual-alone sensory modality there was a significant effect for this variable when the stimuli were presented in the degraded auditory and degraded audiovisual modalities (see section under clear speech – word type – in Table 1). Hence, for each sensory modality, the conversational and clear speech data obtained for the monosyllables and bisyllables were analyzed separately.

Monosyllabic words. The results obtained for the productions of monosyllabic words are displayed in Figure 2. Data are shown for individual talkers. In each panel, speech intelligibility scores obtained for the tokens of conversational and clear speech are displayed by the dashed and filled bars, respectively. The degraded auditory speech intelligibility scores (displayed in the top panel) for conversational speech ranged from 3% (talker I) to 15% (talker A). The degraded auditory speech intelligibility scores for the tokens of clear speech ranged from 5% (talkers E and H) to 26% (talker F). A two-way ANOVA for repeated measures revealed a significant *manner* and *talker* effect as well as a significant *talker* × *manner* interaction (see section under clear speech – monosyllables – in Table 1). Paired two-tailed t tests indicated a significant degraded auditory clear speech advantage for talker F ($t = 7.42$, $df = 11$, $p < .001$), talker I ($t = 2.03$, $df = 11$, $p < .05$), and talker J ($t = 2.55$, $df = 11$, $p < .05$).

The visual speech intelligibility scores are displayed in the middle panel of Figure 2. Visual speech intelligibility scores for conversational speech ranged from 63% (talker E) to 84% (talker G). The visual speech intelligibility scores for the tokens of clear speech ranged from 67% (talker C) to 83% (talker H). A two-way ANOVA for repeated measures revealed that there was not a significant main effect for *manner* of speech. However, there was a significant *talker* effect and a significant *manner* × *talker* interaction (see section under clear speech –

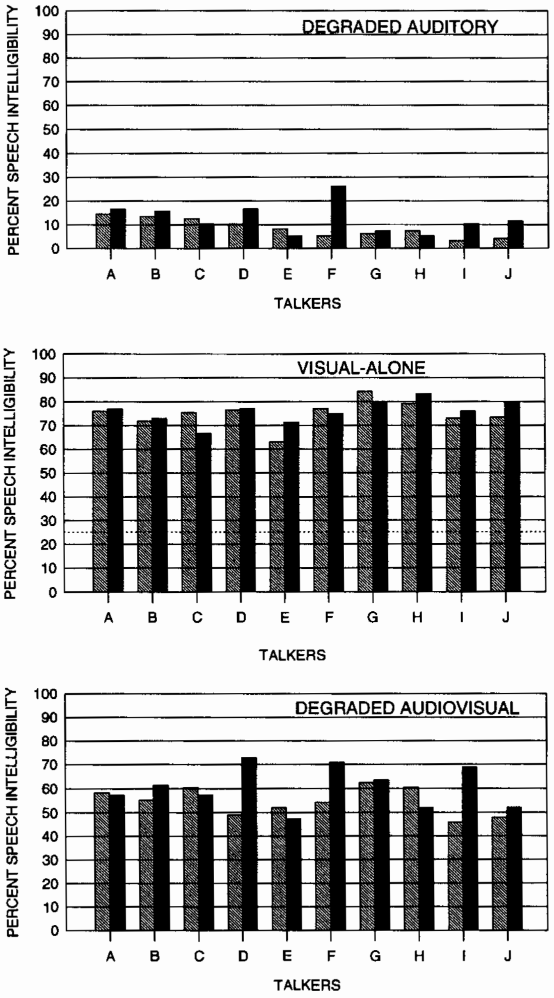

Figure 2. Percent speech intelligibility scores for monosyllabic words. Data are shown for 10 female talkers, in three different sensory modalities: degraded auditory (top panel), visual-alone (middle panel), and degraded audiovisual (bottom panel). For each talker, speech intelligibility scores are shown for tokens of conversational speech (dashed bars) and clear speech (filled bars). In each sensory modality, the speech intelligibility scores are based on data obtained from a different group of subjects. The horizontal dashed line displayed in the middle panel represents the level of chance performance expected for the data obtained under the visual-alone condition.

monosyllables – in Table 1). Paired two-tailed *t* tests indicated a significant visual clear speech advantage for talker E ($t = 3.11$, $df = 23$, $p < .005$) and talker J ($t = -2.29$, $df = 23$, $p < .05$). Talker C obtained a significantly higher visual speech intelligibility score for the tokens of conversational speech than for the tokens of clear speech ($t = 2.38$, $df = 22$, $p < .05$).

The degraded audiovisual speech intelligibility scores are displayed in the bottom panel of Figure 2. The degraded audiovisual speech intelligibility scores for conversational speech ranged from 46% (talker I) to 63% (talker G). The degraded audiovisual speech intelligibility scores for the tokens of clear speech ranged from 48% (talker E) to 73% (talker D). A two-way ANOVA for repeated measures revealed a significant *manner* and *talker* effect as well as a significant *talker* × *manner* interaction (see section under clear speech – monosyllables, in Table 1). Paired two-tailed *t* tests indicated a significant degraded audiovisual clear speech advantage for talker D ($t = 5.06$, $df = 11$, $p < .001$), talker F ($t = 3.37$, $df = 11$, $p < .01$), and talker I ($t = 3.53$, $df = 11$, $p < .005$). In addition, post-hoc tests revealed that talker H obtained a significantly higher degraded audiovisual speech intelligibility score for the tokens of conversational speech than for the tokens of clear speech ($t = 2.97$, $df = 11$, $p < .05$).

Bisyllabic words. The results obtained for the productions of bisyllabic words are displayed in Figure 3. The degraded auditory speech intelligibility scores (displayed in the top panel) for conversational speech ranged from 17% (talker J) to 66% (talker A). The degraded auditory speech intelligibility scores for the tokens of clear speech ranged from 30% (talker J) to 64% (talker F). A two-way ANOVA for repeated measures revealed a significant *manner* and *talker* effect as well as a significant *talker* × *manner* interaction (see section under clear speech – bisyllables – in Table 1). Paired two-tailed *t* tests indicated a significant degraded auditory clear speech advantage for talker E ($t = -2.56$, $df = 11$, $p < .05$), talker F ($t = 4.75$, $df = 11$, $p < .001$), talker G ($t = 3.03$, $df = 11$, $p < .001$), talker I ($t = -3.80$, $df = 11$, $p < .005$), and talker J ($t = -2.80$, $df = 11$, $p < .05$).

The visual speech intelligibility scores are displayed in the middle panel of Figure 3. Visual speech intelligibility scores for conversational speech ranged from 60% (talkers E and J) to 83% (talker G). The visual speech intelligibility scores for the tokens of clear speech ranged from 64% (talker E) to 86% (talker D). A two-way ANOVA for repeated measures revealed a significant *manner* and *talker* effect as well as a significant *talker* × *manner* interaction (see section under clear speech – bisyllables – in Table 1). Paired two-tailed *t* tests indicated a significant visual clear speech advantage for talker A ($t = -3.49$, $df = 23$, $p < .05$), talker I ($t = -2.22$, $df = 23$, $p < .05$) and talker J ($t = 4.66$, $df = 23$, $p < .001$).

The degraded audiovisual speech intelligibility scores are displayed in the bottom panel of Figure 3. The degraded audiovisual speech intelligibility scores for conversational speech ranged from 64% (talker J) to 93% (talker C). The degraded audiovisual speech intelligibility scores for the tokens of clear speech

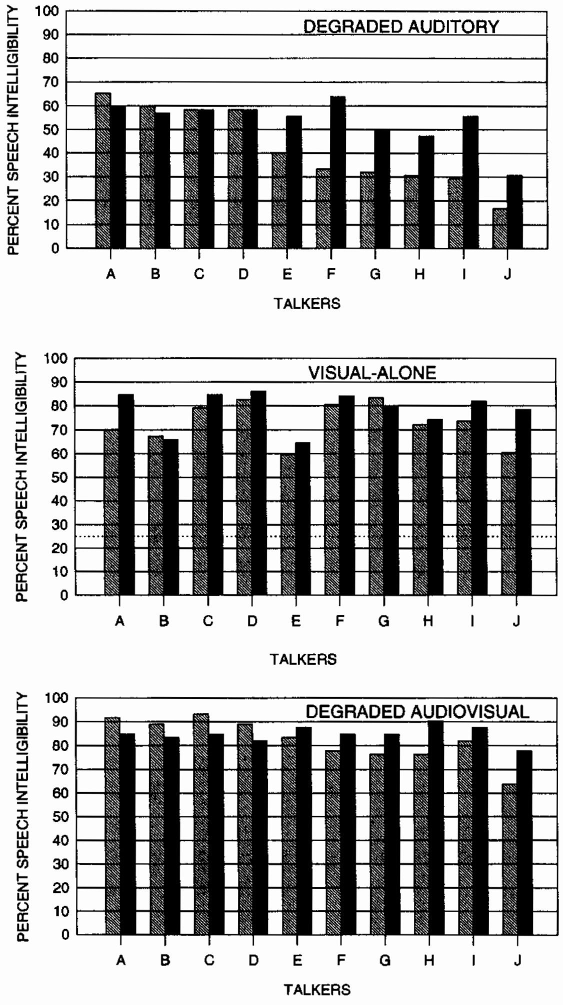

Figure 3. Percent speech intelligibility scores for bisyllabic words. Data are shown for 10 female talkers, in three different sensory modalities: degraded auditory (top panel), visual-alone (middle panel), and degraded audiovisual (bottom panel). For each talker, speech intelligibility scores are shown for tokens of conversational speech (dashed bars) and clear speech (filled bars). In each sensory modality, the speech intelligibility scores are based on data obtained from a different group of subjects. The horizontal dashed line displayed in the middle panel represents the level of chance performance expected for the data obtained under the visual-alone condition.

ranged from 78% (talker J) to 90% (talker H). A two-way ANOVA for repeated measures revealed that there was not a significant main effect for *manner* of speech. However, there was a significant *talker* effect and a significant *talker* × *manner* interaction (see section under clear speech – bisyllables – in Table 1). Paired two-tailed t tests indicated a significant degraded audiovisual clear speech advantage for talker H ($t = -2.06$, $df = 11$, $p < .05$) and talker J ($t = -1.89$, $df = 11$, $p < .05$). It should be noted that for the bisyllabic words none of the talkers produced a significantly higher speech intelligibility score for the tokens of conversational speech than for the tokens of clear speech in any of the sensory modalities under investigation.

Comparison of Clear Speech Effects Across Sensory Modalities

A *clear speech effect* can be defined as the difference between a talker's speech intelligibility for tokens of conversational and clear speech. Statistical analyses of the clear speech effects observed across sensory modalities were not performed because the speech intelligibility scores were not obtained in the same manner (e.g., a closed-set response format was used when the stimuli were presented in the visual-only experimental condition whereas an open-set format was used for the degraded auditory and the degraded audiovisual conditions; different subjects completed the task in each sensory modality). Nevertheless, an informal examination of the clear speech effects displayed by the talkers, across sensory modalities, yielded some noteworthy observations. The clear speech effects observed for each talker in each sensory modality are shown in Figure 4. The results for the monosyllabic and bisyllabic words are displayed in the top and bottom panels, respectively. Overall, positive clear speech effects were observed for each experimental condition (types of words and sensory modality). For the monosyllabic words, the largest overall mean clear speech effect was observed in the degraded audiovisual sensory modality ($M_{av} = 6.66$), followed by the degraded auditory ($M_a = 3.96$) and the visual alone ($M_v = 1.04$) modalities. For the bisyllabic words, the largest overall mean clear speech effect was observed in the degraded auditory sensory modality ($M_a = 11.25$), followed by the visual alone ($M_v = 5.55$), and the degraded audiovisual ($M_{av} = 2.49$) modalities, respectively. The results obtained for individual talkers do not suggest any direct association in the speech intelligibility scores obtained across sensory modalities. Only one talker displayed a significant positive clear speech effect in all three sensory modalities (talker J, for the bisyllabic words). Most of the talkers displayed a positive clear speech effect in some but not all sensory modalities (e.g., talkers D, E, F, I, and J for the monosyllabic words; talkers A, E, F, G, H, and I for the bisyllabic words). The remaining talkers did not produce a significant positive clear speech effect in any of the sensory modalities (e.g., talkers A, B, C, G, and H for the monosyllabic words; talkers B, C, and D for the bisyllabic words). Two talkers produced a significant negative clear speech effect in one sensory modality and no effect in the other sensory modalities (talkers C and H for the monosyllabic words).

An examination of the data displayed in Figures 2 and 3 suggests that the talkers with lower speech intelligibility scores for the recorded tokens of conversational speech displayed the most amount of positive clear speech effects (e.g., for

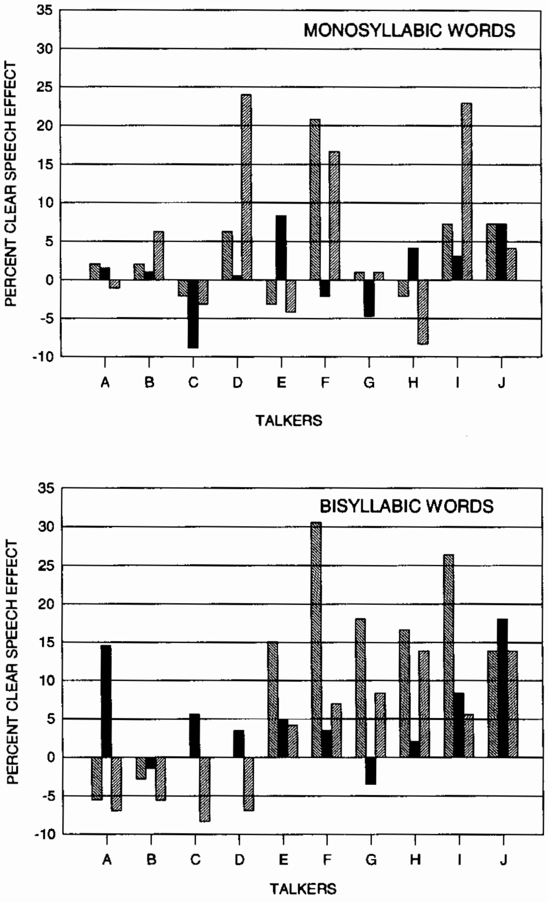

Figure 4. Percent clear speech effects (intelligibility scores for tokens of conversational speech subtracted from intelligibility scores for tokens of clear speech) displayed by individual talkers. Data are shown separately for monosyllabic words (top panel) and bisyllabic words (bottom panel). For each subject clear speech effects are shown for three different sensory modalities: degraded auditory (left-most bar), visual-alone (middle bar), and degraded audiovisual (right-most bar). In each sensory modality, the speech intelligibility scores are based on data obtained from a different group of subjects.

the monosyllabic words see: talker E in the visual-alone condition, and talkers D and I for the degraded audiovisual condition; for the bisyllables see: talker J for the degraded auditory condition, talker J for the visual-alone condition, and talkers H and J for the audiovisual condition). The subjects with relatively high speech intelligibility for the tokens of conversational speech tended to display either no effect or a negative clear speech effect (e.g., for the monosyllabic words see: talker G for the visual-alone condition, talkers H and G for the degraded audiovisual condition; for the bisyllabic words see: talker G for the visual-alone condition, and talkers A, B, C, and D for the degraded audiovisual condition).

DISCUSSION

Overall, in each sensory modality, the perceptual data revealed that speech intelligibility scores were higher for the production of bisyllabic words than for the monosyllables. This finding may be attributable to the nature of the stimuli used in the present investigation. Previous investigations of auditory and visual speech perception have reported that, in general, speech perception performance is better for bisyllabic than monosyllabic words (e.g., Berger, 1972; Davis, 1970). Nevertheless, the range of speech intelligibility scores and the clear speech effects observed, in each of the sensory modalities investigated in the present investigation, were similar for the monosyllabic and the bisyllabic words.

Conversational Speech Intelligibility

The overall mean speech intelligibility scores (for the group of talkers) was substantially different across sensory modalities (M_A = 25.45%; M_V = 74.0%; M_{AV} = 68.40%). These results are most likely due to the nature of the perceptual processes involved in the different tasks, the different variables manipulated across experimental conditions (e.g., the HELOS settings used for the degraded audio and audiovisual conditions), and differences in the response format used across sensory modalities (e.g., closed-set vs. open-set). Notwithstanding the differences in the absolute mean speech intelligibility scores, the range of speech intelligibility scores observed in the present investigation was similar to the results of previous investigations of degraded auditory speech intelligibility (e.g., Cox et al., 1987; Picheny et al., 1985), visual speech intelligibility (e.g., Lesner & Kricos, 1981; Stoker & French-St. George, 1987), and degraded audiovisual speech intelligibility (e.g., Grant & Braida, 1991).

In the present investigation, within each sensory modality, there was a significant difference in speech intelligibility across talkers. These results indicate that, in each sensory modality, the articulation patterns of some talkers yielded significantly better speech recognition scores than those of other talkers. Statistical comparisons of speech intelligibility scores across sensory modalities are not appropriate because the subjects, experimental paradigm, and response format varied across sensory modalities. Nevertheless, a comparison of the relative ranking of the talkers, across sensory modalities, suggests that there is not a direct relationship between a talker's speech intelligibility in one sensory modality and her speech intelligibility in another sensory modality. Some talkers maintained their relative standing in all three sensory modalities. However, for some talkers there was a considerable change in their relative ranking of speech intelligibility scores between the degraded auditory and the visual-alone speech intelligibility scores. These results suggest that articulation patterns that yield highly intelligible speech perception scores in one sensory modality may not have the same positive effects on speech perception in another sensory modality. This finding may have important implications. For example, the data suggest that speech intelligibility scores based on unisensory perceptual data (usually auditory) may not appropriately describe a talker's speech intelligibility during face-to-face conversations. This is particularly relevant for individuals with a hearing loss who rely on visual cues to complement the incomplete and distorted signal they perceive auditorily.

Clinically, the present results suggest that evaluation procedures designed to assess the speech intelligibility of clients with impaired articulation patterns (e.g., individuals with a significant congenital hearing loss or normal-hearing persons with articulation disorders) may yield different results depending on whether auditory-alone or audiovisual speech evaluation procedures are used to assess the client's speech intelligibility. In one study, the speech production competency of 3 children with a profound hearing loss was assessed with stimuli that consisted of syllables spoken in isolation and sentences (Kricos, Lesner, & Lazarus, 1990). Five teachers of children with hearing impairment judged the speech productions under two different conditions: audio-alone and audiovisually. The results revealed some differences in the judgements of speech accuracy between the two scoring conditions. Moreover, the authors reported that the differences in scoring were not consistent and they noted that there appeared to be an interaction among talkers, judges, and sensory modality in which the assessment was completed. The present findings on conversational speech intelligibility, in three different sensory modalities, corroborate the notion that there is an interaction between the talker and the sensory modality in which speech intelligibility is assessed.

Comparison of Conversational and Clear Speech

Overall (for the group of talkers), in each sensory modality, the speech intelligibility scores were slightly better for the tokens of clear speech than for the tokens of conversational speech. However, this positive clear speech effect failed to reach statistical significance with the monosyllabic words presented in a visual-alone sensory modality. The clear speech effects observed in the present investigation are revealing especially because it is very likely that the experimental procedures used in the study did not yield conversational speech intelligibility scores that were truly representative of the talker's speech intelligibility during everyday conversations. First, a talker's conversational speech intelligibility may be inflated when the data are based on the production of words uttered in isolation rather than words excised from samples of running speech. Second, it is possible that the speech utterances produced by the talkers were influenced by the presence of the video camera. Although the talkers were not informed of the specific purpose of the investigation, they were aware that they were being videotaped and they probably surmised that the recordings would be used to investigate aspects of their speech productions. Given those circumstances, it is likely that the talkers were more attentive to the manner in which they produced speech even during the recording of what was operationally defined as tokens of conversational speech. Consequently, the present data most likely represent overestimates of the talkers' intelligibility for conversational speech. The presence of the camera during the recordings of the clear speech tokens would be less deleterious because by definition clear speech is speech that is produced with a conscious attempt to optimize one's speech intelligibility. The clear speech effects (i.e., the difference between clear and conversational speech) observed in the present investigation probably underestimate the effects of clear speech on speech perception during everyday conversations.

An issue of great interest in the present investigation was the clear speech effect produced by individual talkers, both within and across sensory modalities. In each experimental condition, some talkers produced a significant positive clear speech effect. Recall that a specific purpose of the present study was to observe the perceptual effects of clear speech produced by naive talkers. In the present investigation, the talkers were not given any specific instructions, practice, or feedback concerning how to produce clear speech. The results suggest that, in each sensory modality, some talkers had an inherent knowledge of the characteristics of articulatory patterns that would make them more intelligible when they produced clear speech. There is a need for further investigations on the physical attributes of clear speech. Specifically, acoustic and kinematic analyses of the speech patterns (both conversational and clear speech utterances) of talkers who produced significant clear speech effects may reveal some characteristics of articulation patterns that improve speech intelligibility in a given sensory modality.

For the monosyllabic words, some talkers produced a *negative* clear speech effect (i.e., the speech intelligibility scores were higher for the tokens of conversational speech than for the tokens of clear speech). Specifically, talker C produced a significant negative clear speech effect in the visual-alone and talker H produced a significant negative clear speech effect in the degraded audiovisual condition. Also, it should be noted that in each sensory modality, the majority of talkers produced neither a significant positive nor a significant negative clear speech effect. These findings suggest that many individuals may benefit from instructions and training on the production of articulatory patterns that could enhance their speech intelligibility.

The magnitude of positive clear speech effects displayed by some talkers (in each sensory modality) suggests that individuals with hearing impairment could benefit substantially from communication partners who use clear speech. For example, in the degraded auditory condition, some talkers displayed improvements that exceeded 20% for the monosyllabic words and 30% for the bisyllabic words for the tokens of clear speech. These results are consistent with previous investigations of auditory-clear speech effects that have been reported for syllables (e.g., Uchanski, 1988) and sentences (e.g., Picheny et al., 1985; Uchanski, 1988).

The present results indicate that some individuals do not inherently know how to produce clear speech (or worse, their attempts to produce clear speech resulted in poorer speech perception scores in one or more sensory modality). Those individuals could benefit from instructions or training on how to produce clear speech. Also, individuals with hearing impairment should be aware of the characteristics of clear speech so that they can request those articulation patterns from their communication partners. However, it is likely that more research on the perceptual effects and the physical characteristics of clear speech will be required before the use of clear speech can be fully exploited clinically. For example, the effects of clear speech on speech perception in different sensory modalities should be investigated with sentence-length stimuli. Those results would be more representative of the effects of clear speech during everyday conversations.

An interesting aspect of the present investigation was that the results made it possible to examine (albeit indirectly) the effects of clear speech on speech perception in different sensory modalities. The present findings suggest that the articulatory patterns that contribute to improvements in speech perception in one sensory modality may not have any effect (or may actually have a deleterious effect) on speech perception in other sensory modalities (e.g., talkers who produced a positive auditory clear speech effect but a negative visual clear speech effect). Yet some talkers displayed positive clear speech effects in more than one sensory modality. The latter finding indicates that in some cases it is possible to modify the manner in which speech is produced so that beneficial effects will be observed in all three sensory modalities. Additional research is required to identify the articulatory patterns that result in improvements in speech perception in all sensory modalities.

Finally, the present findings warrant a comment on the use of the term *clear speech.* In the present investigation, the term clear speech was used to describe the speech intelligibility scores obtained for tokens of speech generated following a request to produce speech patterns that would make the talker more intelligible (i.e., "articulate each word clearly as if you were talking to someone who had difficulty understanding what you are saying."). This definition of the term clear speech is consistent with the definitions used by previous investigators (e.g., Moon & Lindblom, 1989; Picheny et al., 1985). It defines clear speech according to the demands of a communicative situation rather than the outcome of some measure of speech intelligibility per se. Consequently, as was observed in the present investigation, it is possible to encounter the awkward situation whereby a talker's production of clear speech results in speech intelligibility scores that are lower than the person's speech intelligibility scores for tokens of conversational speech. Perhaps one way to eliminate this type of ambiguity would be to specify whether a given speech intelligibility score represents a measure of *solicited* or *measured* clear speech.

SUMMARY AND CONCLUSIONS

The present study investigated various aspects of speech intelligibility (operationally defined as the effects of speech tokens produced by 10 female talkers on the speech perception performances obtained from three different groups of normal-hearing adults). The results confirmed that significant differences in degraded auditory, visual, and degraded audiovisual speech intelligibility exist across talkers. A comparison of tokens of conversational and clear speech obtained from each talker revealed that most of the 10 talkers were unable to produce a significant clear speech effect spontaneously. Additional studies are required to further elucidate the effects of clear speech on speech perception in different sensory modalities.

ACKNOWLEDGEMENTS

Components of the research reported in the present manuscript were presented at the Summer Institute of the Academy of Rehabilitative Audiology in Austin, Texas: June, 1992.

This work was supported by a Health System-Linked Research Unit Grant (Ontario Ministry of Health) to the Hearing Health Care Research Unit (HHCRU: Department of Communicative Disorders, The University of Western Ontario) and by a research grant from the Natural Science and Engineering Research Council of Canada awarded to the first author (J-PG, École d'orthophonie et d'audiologie, Université de Montréal). Additional support was received from the "Comité d'attribution des fonds de recherche" (CAFIR: Université de Montréal) and by the "Fonds de développement à la recherche" (FDR: Université de Montréal) awarded to the first author (J-PG, École d'orthophonie et d'audiologie, Université de Montréal). The authors are grateful to the Editor and three external Reviewers for their comments on earlier drafts of this manuscript.

REFERENCES

American National Standards Institute. (1969). *Specifications for audiometers* (ANSI S3.6-1969). New York: Author.

Berger, K.W. (1972). *Speechreading: Principles and methods.* Baltimore: National Educational Press.

Bilida, N.E. (1992). *Effects of clear speech on auditory-alone and auditory-visual speech perception.* Unpublished M.Sc. research project, The University of Western Ontario, London, Ontario, Canada.

Boothroyd, A. (1985). Measurement of speech production in hearing impaired children: Some benefits of forced-choice setting. *Journal of Speech and Hearing Research, 28,* 185-196.

Boothroyd, A. (1988). Linguistic factors in speechreading. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review, 90* (5), 77-87.

Carney, A. (1986). Understanding speech intelligibility in the hearing-impaired. *Topics in Language Disorders, 6* (3), 47-59.

Chen, F.R. (1980). *Acoustic characteristics and intelligibility of clear speech at the segmental level.* Unpublished master's thesis, MIT, Cambridge. (Cited in Uchanski, 1988).

Cox, R.M., Alexander, G.C., & Gilmore, C. (1987). Intelligibility of average talkers in typical listening environments. *Journal of the Acoustical Society of America, 81,* 1598-1608.

Creelman, C.D. (1957). Case of the unknown talker. *Journal of the Acoustical Society of America, 29,* 655.

Davis, H. (1970). Audiometry: Pure tone and simple speech tests. In H. Davis & R.S. Silverman (Eds.), *Hearing and deafness* (3rd ed.) (pp. 179-220). New York: Holt, Rinehart and Winston.

Davis, H., & Silverman, R.S. (1970). *Hearing and deafness* (3rd ed.). New York: Holt, Rinehart and Winston.

De Filippo, C.L. (1991). Good talker/poor talker: What makes the difference? *Contact, 6* (3), 30-32.

Erber, N.P. (1988). *Communication therapy for hearing-impaired adults.* Abbotsford, Victoria, Australia: Clavis Publishing.

Erber, N.P. (1993). *Communication and hearing loss.* Melbourne, Victoria, Australia: Clavis Publishing.

Franks, R.J. (1979, December/January). The influence of exaggerated mouth movements on lipreading. *Audiology & Hearing Education,* 12-16.

Gagné, J.-P., & Erber, N.P. (1987). Simulation of sensorineural hearing loss. *Ear and Hearing, 8,* 232-243.

Gagné, J.-P., Tugby, K., & Michaud, J. (1991). Development of a Speechreading Test on the Utilization of Contextual Cues (STUCC): Preliminary findings with normal-hearing subjects. *Journal of the Academy of Rehabilitative Audiology, 24,* 157-170.

Grant, K.W., & Braida, L.D. (1991). Evaluating the articulation index for auditory-visual input. *Journal of the Acoustical Society of America, 89,* 2952-2960.

Hood, J.D., & Poole, J.P. (1980). Influence of the speaker and other factors affecting speech intelligibility. *Audiology, 19,* 434-455.

Jackson, P. (1988). The theoretical minimal unit for visual speech perception: Visemes and coarticulation. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review, 90* (5), 99-115.

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

Kaplan, H., Bally, S.J., & Garretson, C. (1985). *Speechreading: A way to improve understanding.* Washington, DC: Gallaudet College Press.

Kreul, E.J., Bell, D.W., & Nixon, J.C. (1969). Factors affecting speech discrimination test difficulty. *Journal of Speech and Hearing Research, 12,* 281-287.

Kricos, P.B., & Lesner, S.A. (1982). Differences in visual intelligibility across talkers. *The Volta Review, 84,* 219-225.

Kricos, P., Lesner, S., & Lazarus, G. (1990). Influence of visual information on speech assessment with hearing-impaired children. *The Volta Review, 92,* 213-222.

Lesner, S.A., & Kricos, P.B. (1981). Visual vowel and diphthong perception across speakers. *Journal of the Academy of Rehabilitative Audiology, 14,* 252-258.

Moon, S.J., & Lindblom, B. (1989). Formant undershoot in clear and citation-form speech: A second progress report. *Quarterly Progress Status Report, 1,* 121-123. Stockholm: Department of Speech Communication.

Mullennix, J.W., Pisoni, D.B., & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85,* 365-378.

Palmer, J.M. (1955). The effect of speaker differences on the intelligibility of phonetically balanced word lists. *Journal of Speech and Hearing Disorders, 20,* 192-195.

Payton, K.L., Uchanski, R.M., & Braida, L.D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America, 95,* 1581-1592.

Penrod, J.P. (1979). Talker effects on word-discrimination scores of adults with a sensorineural hearing impairment. *Journal of Speech and Hearing Disorders, 44,* 340-349.

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,* 96-103.

Picheny, M.A., Durlach, N.I., & Braida, L.D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research, 29,* 434-446.

Picheny, M.A., Durlach, N.I., & Braida, L.D. (1989). Speaking clearly for the hard of hearing: An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research, 29,* 600-603.

Stoker, R.G., & French-St. George, M. (1987). Lipreadability in normal hearing speakers: Effects of short-term training [Abstract]. *Asha, 29* (10), 77.

Tye-Murray, N., & Schum, L. (1994). Conversation training for frequent communication partners. In J.-P. Gagné & N. Tye-Murray (Eds.), *Research in audiological rehabilitation: Current trends and future directions* [Monograph]. *Journal of the Academy of Rehabilitative Audiology, 27,* 209-222.

Uchanski, R.M. (1988). *Spectral and temporal contributions to speech clarity for hearing impaired listeners.* Unpublished doctoral dissertation, MIT, Cambridge.

Verbrugge, R.R., Strange, W., Shankweiler, D.P., & Edman, T.R. (1976). What information enables a listener to map a talker's vowel space. *Journal of the Acoustical Society of America, 60,* 198-212.

Weinhouse, I., & Miller, M. (1963). Discrimination scores for two lists of phonetically balanced words. *Journal of Auditory Research, 3,* 9-14.

Williams, C.E., & Hecker, M.H. (1968). Relation between intelligibility scores for four methods and three types of speech distortions. *Journal of the Acoustical Society of America, 44,* 1002-1006.

APPENDIX A

TEST WORDS (IN CAPITAL LETTERS) AND RESPONSE FOILS USED FOR THE VISUAL-ONLY WORD RECOGNITION TASK

CHEAT	sheep	teach	cheer
CURSE	hurt	girls	purse
FEED	mead	tweed	fear
GREEK	three	dream	beep
GRUDGE	watch	fudge	great
MAST	bath	math	task
PIPE	bike	dime	bite
TRUE	glue	through	blue
COWBOY	coward	cowbell	playtoy
DAYBREAK	subway	playmate	payday
GRANDSON	gryphon	greatstuff	watchman
GREYHOUND	whitehouse	playground	profound
HEADLIGHT	alright	bedside	nightlife
PANCAKE	padlock	manmade	brandname