Influence of Visual Information on the Intelligibility of Dysarthric Speech

Connie K. Keintz
Kate Bunton
Jeannette D. Hoit
University of Arizona, Tucson

Purpose: To examine the influence of visual information on speech intelligibility for a group of speakers with dysarthria associated with Parkinson's disease.
Method: Eight speakers with Parkinson's disease and dysarthria were recorded while they read sentences. Speakers performed a concurrent manual task to facilitate typical speech production. Twenty listeners (10 experienced and 10 inexperienced) transcribed sentences while watching and listening to videotapes of the speakers (auditory-visual mode) and while only listening to the speakers (auditory-only mode).
Results: Significant main effects were found for both presentation mode and speaker. Auditory-visual scores were significantly higher than auditory-only scores for the 3 speakers with the lowest intelligibility scores. No significant difference was found between the 2 listener groups.
Conclusions: The findings suggest that clinicians should consider both auditory-visual and auditory-only intelligibility measures in speakers with Parkinson's disease to determine the most effective strategies aimed at evaluation and treatment of speech intelligibility decrements.

Key Words: Parkinson's disease, auditory-visual cues, speech perception, listener experience

Speech intelligibility is of paramount concern in both the evaluation and management of dysarthria. Individuals with dysarthria often exhibit reduced intelligibility, defined here as "the degree to which a speaker's message can be recovered by a listener" (Kent, Weismer, Kent, & Rosenbek, 1989, p. 483). Current clinical methods for measuring speech intelligibility reflect the interaction between a speaker and a listener under given communication conditions (Ansel, 1985). Years of research have shown that there are many factors that can influence intelligibility measures. These include (a) severity of the intelligibility impairment (Yorkston & Beukelman, 1978), (b) speech rate (Yorkston & Beukelman, 1981), (c) type of speech stimulus, (d) scoring method (Beukelman & Yorkston, 1980), (e) predictability of stimuli (Duffy & Giolas, 1974), (f) listener familiarity with the speech sample (Beukelman & Yorkston, 1980), and (g) listener experience with the speech disorder and/or the individual speaker (e.g., Beukelman & Yorkston, 1980; Platt, Andrews, Young, & Quinn, 1980; Yorkston & Beukelman, 1980). In addition, acoustic and nonacoustic cues produced by the speaker have been shown to influence measures of speech intelligibility.

A listener who is attempting to understand a speaker's verbal message relies on two types of information (Lindblom, 1990). One is signal-dependent, and one is signal-independent. Information taken entirely from the acoustic signal is described as signal-dependent. In speakers with dysarthria, compromised signal-dependent information can contribute to decreased intelligibility.

Information other than that provided by the acoustic signal is referred to as signal-independent. Signal-independent information includes the listener's knowledge of the language and aspects of the communicative context. Hustad, Jones, and Dailey (2003) have described three types of communicative knowledge that may be used by listeners to decode a spoken message. They include (a) linguistic knowledge, which defines a listener's expectations for semantics, syntax, and phonology; (b) paralinguistic knowledge, such as that related to gestures, facial expression, and speech-related movements; and (c) experiential knowledge, which refers to shared knowledge of culture and experiences between the listener and speaker. Attention to signal-independent information can help a listener understand a spoken message when the signal-dependent information is degraded.

Interactions between signal-dependent and signal-independent information may be important for listeners attempting to understand the message of speakers with impaired speech intelligibility. Lindblom has proposed a

1058-0360/07/1603-0222

model of mutuality (1990) that describes the interaction between signal-dependent and signal-independent information in such cases. Lindblom suggested that because speakers with intelligibility impairments provide listeners with reduced signal-dependent (acoustic-phonetic) information, listeners may be more reliant on the signal-independent (linguistic-contextual) knowledge. That is, signal-independent information can be used to fill in the gaps left by incomplete or compromised signal-dependent information. Lindblom proposed that these two sources of information are negatively correlated. This means that when signal-dependent information (the speech signal) is adequate, the listener should be able to understand the message in the absence of signal-independent information. By contrast, when the signal-dependent information is inadequate, signal-independent information may be crucial to understanding a speaker's message.

Because speech intelligibility is defined as the amount of speech understood from the acoustic signal alone, most current clinical measurement tools for speech intelligibility allow listeners access to only signal-dependent information (audio-recordings) and limit listener access to signal-independent information. Of specific interest here is listener access to visual information. The most common measurement protocol for intelligibility involves audiotape-recording the speaker and then asking the listener to transcribe words and/or sentences (Enderby, 1983; Tikofsky & Tikofsky, 1964; Yorkston & Beukelman, 1980; Yorkston, Beukelman, & Tice, 1996) or use a multiple-choice selection procedure (Yorkston & Beukelman, 1980). The resultant intelligibility score is the percentage of words correctly transcribed or selected. There are two measurement tools that also provide information concerning phonetic errors (Kent et al., 1989; Platt et al., 1980). Only one published measurement tool, the Frenchay Dysarthria Assessment (Enderby, 1983), allows (but does not require) the clinician to see the client's face.

There is considerable evidence that visual information influences intelligibility (for reviews, see Massaro, 1987; Summerfield, 1987). For example, when the fidelity of normal speech is degraded by background noise, visual information provided by a speaker's face can enhance speech intelligibility (Neely, 1956; O'Neill, 1954; Sumby & Pollack, 1954).

Visual information can also influence intelligibility of speakers with disordered speech. In studies of hearing-impaired speakers, intelligibility scores were higher under conditions in which listeners had access to both auditory and visual information compared to conditions in which they had access only to auditory information (Menke, Oschner, & Testut, 1983; Monsen, 1983; Siegenthaler & Gruber, 1969). Similar findings have been reported for speakers with laryngectomies using esophageal speech (Berry & Knight, 1975; Hubbard & Kushner, 1980).

The nature of the influence of visual information on the intelligibility of speakers with dysarthria is less clear. Barkmeier (1988) examined 12 speakers with dysarthria resulting from various etiologies. Results indicated higher intelligibility scores when listeners (10 experienced and 10 inexperienced) watched and listened to videotapes of these speakers than when they only listened to them. It was not possible to determine the effects of condition on intelligibility of individual speakers, as only group means were reported. However, there was an order effect. Specifically, when the auditory-only condition came first, intelligibility scores were significantly higher in the second (auditory-visual) session. By contrast, when the auditory-visual condition came first, intelligibility scores were not significantly higher than the auditory-only session. Results of this study also indicated that scores obtained from experienced listeners were significantly higher than those obtained from inexperienced listeners.

Hunter, Pring, and Martin (1991) examined influences of auditory-only and auditory-visual presentation modes in 8 speakers with dysarthria related to cerebral palsy. Results showed that the speakers with moderate dysarthria were more intelligible in the auditory-visual condition than in the auditory-only condition. By contrast, the speakers with severe dysarthria had similar intelligibility scores in the two presentation modes. The scores obtained from 16 experienced and 16 inexperienced listeners did not differ significantly. It should be noted that the severity levels of these speakers were determined by gross clinician perceptual ratings, rather than by use of an intelligibility score.

Garcia and Cannito (1996) studied a single speaker with severe flaccid dysarthria. Stimuli included two types of utterances, those that were high in predictability and those that were low in predictability, both of which were produced with and without gestures. Stimuli were presented to listeners in three modes: auditory-visual, auditory-only, and visual-only. Results for the low predictive utterances in the ungestured condition (the condition most closely resembling the conditions of the present study) showed no significant differences between auditory-visual and auditory-only conditions for 48 inexperienced listeners.

Hustad and Cahill (2003) reported mixed findings in a study of auditory-only versus auditory-visual modes of presentation for a group of 5 speakers with dysarthria associated with cerebral palsy. The listeners were 100 college students with little to no experience listening to someone with a communication disorder. Results showed that only 1 of the 5 speakers demonstrated significantly higher intelligibility scores in the auditory-visual condition than in the auditory-only condition.

It is difficult to determine why the findings varied so much in these studies of the influence of visual information on the intelligibility of dysarthric speech. It may be that factors related to speaker characteristics could account for some of the differences across studies. The speakers who have been studied were heterogeneous, representing a variety of etiologies, a range of severity, and several types of dysarthria.

Other factors that may have contributed to differences in results across these studies of intelligibility of dysarthric speech relate to the degree to which semantic/syntactic predictability was taken into account in the development of the stimuli. Specifically, only Hunter et al. (1991) and Garcia and Cannito (1996) used stimuli that were balanced for semantic predictability (i.e., predictability of words within a sentence context) across presentation mode (i.e., auditory-visual and auditory only), even though semantic predictability has been shown to influence intelligibility (Kalikow, Stevens, & Elliott, 1977). Further, only Hustad and Cahill (2003) used sentences that were syntactically predictable (subject-verb-object).

Another important factor that was not taken into account by any of the studies cited here (Barkmeier, 1988; Garcia & Cannito, 1996; Hunter et al., 1991; Hustad & Cahill, 2003) relates to visual information associated with the speech sample. The amount of visual information associated with different sounds varies substantially. Certain sounds, such as bilabial consonants (/b/, /p/, or /m/) or lip-rounded vowels (/u/), provide a great deal of visual information to listeners, whereas others, such as velar consonants (/g/, /k/, or /ng/) or the glottal fricative /h/, provide little or no visual information (for further discussion, see Lidestam & Beskow, 2006). Fortunately, it is possible to create sets of speech stimuli that are balanced for visual information, as well as auditory information, so as to control for this variable (MacLeod & Summerfield, 1990; Rosen & Corcoran, 1982; Rosenblum, Johnson, & Saldana, 1996).

The present study examined the influence of visual information on intelligibility for a group of speakers with dysarthria representing a single etiology and using stimuli that controlled for semantic predictability, syntactic predictability, and visual information. The etiology chosen was Parkinson's disease, because speakers with Parkinson's disease generally exhibit characteristics that affect visual information associated with speech, including a masked face, reduced amplitude of movement during speech production (e.g., lips, jaw), and an accelerated speech rate (Duffy, 2005). The following questions were addressed as they pertain to speech associated with Parkinson's disease:

1. Does the presentation mode (auditory-only vs. auditory-visual) influence speech intelligibility of adults with dysarthria associated with Parkinson's disease?
2. Does experience of the listener (experienced vs. inexperienced with dysarthric speech) influence intelligibility?
3. Is there an interaction between presentation mode and listener experience?

Method

Participants

Three groups of participants were included in this study: speakers, inexperienced listeners, and experienced listeners. All participants signed consent forms that had been approved by the institutional review board at the University of Arizona.

Speakers

Eight participants, 6 men and 2 women with Parkinson's disease and associated dysarthria, served as speakers (see Table 1). They ranged in age from 57 to 82 years, had English as a first language, reported normal to corrected-to-normal vision, and did not have a beard, mustache, or facial deformity that may have impeded visualization of their facial movements. Speakers were required to pass a pure-tone hearing screening at 40 dB HL for octave frequencies of 500, 1000, 2000, and 4000 Hz in at least one ear (Grason-Stadler GSI 17 audiometer). This threshold is considered typical for individuals over the age of 50, and hearing loss below this level is unlikely to affect speech production (Morrell, Gordon-Salant, Pearson, Brant, & Fozard, 1996). All speakers exhibited intelligibility impairments as judged by at least one investigator and as demonstrated by a score of less than 95% on the Sentence Intelligibility Test (Yorkston et al., 1996) based on responses of an unfamiliar listener.

Five master clinicians, who each reported more than 20 years of clinical experience working with patients with

TABLE 1. Demographic and descriptive information on 8 speakers.

Speaker	Sex	Age	Sentence Intelligibility Test scores	Perceptual speech characteristics
A	M	81	38	Harsh/hoarse voice quality Imprecise consonants Variable rate
B	M	60	42	Harsh/hoarse voice quality Imprecise consonants Variable rate
C	F	79	70	Breathy/hoarse voice quality
D	M	74	75	Harsh/hoarse voice quality Imprecise consonants Variable rate
E	M	72	80	Harsh/hoarse voice quality Imprecise consonants
F	F	57	86	Breathy voice quality Imprecise consonants Variable rate
G	M	82	92	Harsh voice quality Imprecise consonants Variable rate
H	M	73	82	Harsh/hoarse voice quality Imprecise consonants

dysarthria, provided descriptions of each speaker's speech. These clinicians were asked to listen to audio samples of the speakers' vowel prolongations, rapid syllable repetitions, and paragraph reading, and to indicate the presence of deviant perceptual dimensions based on the Darley, Aronson, and Brown (1969a, 1969b) taxonomy. Deviant perceptual features that were noted by at least four of the five master clinicians are shown in Table 1. All speakers were judged to have perceptual deviancies related to voice quality (harsh/hoarse for male speakers and breathy/hoarse for female speakers). Five of the 8 speakers were judged to have variable rate, and 7 of the 8 were judged to have imprecise consonants. No other deviant perceptual characteristics were identified by four of the five judges.

Listeners

Twenty listeners were included in the study. All met the following criteria: (a) had English as their primary language; (b) reported normal or corrected-to-normal vision; (c) passed a pure-tone hearing screening at 20 dB HL at 500, 1000, 2000, 4000, and 6000 Hz in both ears (GSI 17 audiometer; American Speech-Language-Hearing Association, 1997); (d) had no specific knowledge of the study; and (e) had no prior exposure to the speakers who participated in the study.

Listeners were divided into two groups (10 per group), an inexperienced and an experienced group. The inexperienced listeners were recruited from the university student community. They ranged in age from 19 to 24 years, had no prior experience evaluating and/or treating speakers with dysarthria, had no course work in motor speech disorders, and reported no regular (daily or weekly) contact with a speaker with dysarthria. The experienced listeners were selected from a pool of local speech-language pathologists. They ranged in age from 27 to 54 years, were certified by the American Speech-Language-Hearing Association as speech-language pathologists, were currently working with adults with dysarthria, and had more than 6 months of professional experience with this population.

Stimuli Selection

The stimulus sentences that were read aloud by each speaker were developed to take into account semantic predictability, syntactic structure, and auditory and visual characteristics. They consisted of 8 lists of 15 sentences each (see Appendix) based on the Institute of Hearing Research (IHR) Audio-Visual Sentence Lists (MacLeod & Summerfield, 1990). Sentences were derived from lists developed for speech-in-noise tests based on Bench and Bamford's (1979) British sentences. MacLeod and Summerfield adapted these sentences into 10 sets of 15 sentences equalized for lipreading and presence of visibly distinct consonants using procedures described by Rosen and Corcoran (1982). Their 10 lists contained an equal number of sentences with the same syntactic structure (e.g., subject-verb-object).

The semantic predictability of these lists was determined in a separate study in which 90 university students completed three different forms of a fill-in-the-blank test. A different key (high-content) word from each sentence was left as a blank on each form of the sentence. For example, "____ moved the furniture" (Form 1), "They ____ the furniture" (Form 2), and "They moved the _____" (Form 3). The students were asked to write the most likely word to occur in each blank. Each student only received one form of each sentence. Results of statistical analysis (t tests run on the number of words guessed correctly) indicated that 8 of the 10 IHR lists were statistically similar for semantic predictability. These 8 lists could be described as "low predictability" with percentages ranging from 19% to 28%. In summary, Lists 2 through 8 and List 10 taken from the MacLeod and Summerfield (1990) study were selected based on low semantic predictability and similar syntactic structure, and were balanced for amount of visual information. Low predictability stimuli were preferred because listeners would be less able to guess at words they did not understand.

Stimulus-Recording Procedures

A recording session was scheduled if speakers met the initial criteria for study participation (i.e., interest in participation, reduced intelligibility, similar dysarthria characteristics). When possible, this session was conducted in an IAC sound-treated room. Three of the 8 speakers were not able to be recorded in this room due to transportation issues and were instead recorded in a quiet environment with background and lighting similar to that used in the sound-treated room. The sessions were scheduled according to when the speaker felt his or her intelligibility would demonstrate the greatest impairment, with medication cycle and fatigue considered as factors.

Equipment Setup

Each speaker was seated comfortably in a chair and encouraged to place his or her arms on the chair's arms for stability. Two of the 8 speakers sat in their wheelchairs. Each speaker was positioned in front of a neutral-colored background and wore a black cloak to minimize visual distractions. Microphone audio signals were recorded simultaneously onto a digital audiotape (DAT) recorder (Tascam Digital Audio Tape Recorder Model DA-P1) and videotape recorder using a lapel microphone (Audiotechnica AT 8533) attached to the speaker's collar. The microphone did not obscure the speaker's face or mouth. The mouth-to-microphone distance was kept constant for each speaker throughout the recording session, but varied between speakers due to differences in physical size (range = 9–13 cm). Video and audio samples were recorded with a video recorder (Sony Video Camera Recorder Model CCD-TR101) positioned directly in front of each speaker. The camera was set to capture each speaker's face centered from the shoulders up. A constant distance (approximately 6 ft) between speaker and camera was maintained across speakers, with the camera set on auto-focus. The angle of the camera was maintained parallel to the floor, and a vertical movement of the tripod platform was adjusted for height differences in speakers.

Speech Sample Recording

Orthographic representations of stimulus sentences were printed on neutral-colored paper with typed lettering (Times New Roman, 74 point) and positioned on a music stand placed directly above the video camera lens. This placement ensured that the speaker's gaze was just above the camera lens so that the speaker appeared on tape to be looking at the camera. The speaker read the sentences aloud from the cards. The speaker was instructed to speak naturally as though having a conversation. If a speaker made a reading error, that sentence was set aside and presented again at the end of that sentence block. Delaying a repetition of the sentence was judged to minimize performance effects that may have occurred as the speaker attempted to correct the sentence immediately. Planned 2-min breaks were taken after each sentence block (15 sentences). Additional breaks were allowed at any time during the recording session when requested by the speaker. Speakers read 120 sentences (excluding any sentences that required repetition), and the lists and sentences within the lists were randomized for each speaker in order to reduce effects of fatigue. The actual recording of sentences took less than 1 hr.

Recordings from 4 pilot speakers with Parkinson's disease revealed a substantial performance effect in which they produced highly intelligible speech that was not consistent with their typical performance during conversation (as judged by the first two authors and each speaker's spouse or family member). Significant differences between clinical performance and ecological manifestations of dysarthria have been reported previously (Sarno, 1968; Weismer, 1984). To reduce performance effects and elicit speech that was representative of the speaker's typical speech production, a dual-task paradigm such as those previously used in other speech studies was implemented (Dromey & Benson, 2003; Ho, Iansek, & Bradshaw, 2002).

The underlying assumption with a dual-task paradigm is that if attentional capacity is limited, and attention is divided between two tasks performed simultaneously, then performance on one or both will be negatively affected (Kahneman, 1973; Wickens, 1984). In general, studies using dual-task paradigms report that if one of the tasks is novel, complex, or speeded, it will have a more negative effect on the other task than if both tasks are relatively simple (Morris, Iansek, Matyas, & Summers, 1996). Therefore, to ensure that including a secondary task did not have an overly negative effect on speech production and thus create a situation in which the characteristics of the dysarthria were exaggerated, we employed a motor task that was likely familiar to the speaker, was relatively simple, and did not place time-related demands on the speaker. The dual task required that the speaker hold a nut in his or her dominant hand and screw a bolt with the other hand. Speakers performed the manual task under the black cloak so that they could not see the objects or their hands and so that the hand movements could not be seen on the videotape. An investigator watched the speakers from the side to ensure that they continued to screw the bolt while reading the sentences.

Stimulus Set Preparation

Video recordings were transferred to a personal computer via FireWire (Institute of Electrical and Electronics Engineers 1394 interface). Digital video software (Adobe Premiere Pro; Adobe Systems, 2004) was used to edit the video recordings of each speaker. It was found that an automatic gain control device on the video recorder amplified extraneous background noise. To eliminate this noise, audio files from the DAT recordings were matched with the digital audio files from the video recording. A computer software program (MATLAB Version 7.0) was used to align the two audio signals. A cross-correlation was computed between the audio signal from the DAT and videotaped samples using the first 200,000 sample points to determine the time difference between them. The DAT audio signal was then shifted to match the sample from the videotape. To accomplish this, either points were taken away or the beginning of the audio file was padded with extra samples of zero amplitude. Following the alignment of the two audio signals, the higher quality DAT audio files were imported into the Premiere Pro software program. The original audio sample was then deleted, leaving the video file aligned with the DAT audio file.

Each sentence produced by each speaker was dubbed into an auditory-visual movie file and an auditory-only movie file. For each auditory-visual movie file, the image of the speaker appeared for approximately 1 s before the sentence was produced to allow listeners a chance to view each speaker at rest. For each auditory-only movie file, a black screen with a small blue square was presented while the audio portion of the movie was played. A 12-s pause was inserted between each sentence to allow for listeners to transcribe the sentence. A randomized sentence order was used for each speaker and each list. Edited movie files were recorded onto DVD format for presentation.

Listening Tasks

Each of the 20 listeners was assigned to a random presentation order for speakers, sentence lists, and order of presentation mode (auditory-visual or auditory-only first). Listeners were presented with all 8 speakers during each listening session. Listeners completed two sessions, with the second session occurring 7 to 10 days after the first. Although listeners heard the same sets of sentences in these two sessions, they did not hear the same speaker producing the same sentences. All listening sessions were conducted individually in a quiet listening environment with dimmed lighting to allow optimal visualization of the video screen. The listener wore headphones (AKG Model K 240 DF) and was seated at a table approximately 2 ft in front of a 17-in. digital video monitor (positioned at eye level). The output level was set to a comfortable listening level prior to the start of the listening sessions and was maintained throughout all sessions for that listener.

Written and oral instructions (for auditory-visual or auditory-only) were given to listeners at the beginning of the session. Listeners were instructed to write down as accurately as possible what they heard each speaker say.

Listeners were first presented with two practice sentences taken from the two eliminated IHR lists (these were different for each speaker). The experimental sentences were then presented, each one time. In total, each listener transcribed 20 sentences for each speaker in a session, including 2 practice sentences, 3 reliability items, and 15 sentence items for analysis.

Scoring

Intelligibility for each speaker was determined by counting the number of correctly identified words per sentence and dividing by the total number of words possible for each sentence (four to seven words per sentence). The scores from each sentence for each listener were then totaled. A mean score was computed across the 10 listeners in each listening group (experienced and inexperienced). Synonyms or responses reflecting morphological variations, such as *cat* for *cats,* were considered incorrect. Misspellings (e.g., *theif* for *thief*) and homonyms (e.g., *their* for *they're* or *rode* for *rowed*) were accepted as correct. Two raters scored the number of correct words per sentence, and these scores were compared. Less than 5% of these scores were not in agreement, and in these cases, the raters discussed the scores and came to a consensus before a final score was assigned.

Experimental Design and Analysis

A total of 4,800 transcriptions were collected (20 listeners × 15 sentences × 2 modes × 8 speakers). Mean intelligibility scores and standard deviations were computed for each speaker in each presentation mode. Mean scores were used for inferential statistical analyses.

A four-way analysis of variance (ANOVA) was used to examine four main effects: presentation mode (auditory-only vs. auditory-visual), speaker, listener experience (experienced vs. inexperienced), and order of presentation (auditory-only/auditory-visual vs. auditory-visual/auditory-only). Speaker and presentation mode were examined as within-subject variables and listener experience and order as between-subject variables. This ANOVA used an interaction model, with the alpha level set at .05 per family of tests.

Two post hoc tests were conducted. The first post hoc analysis examined differences within and across presentation mode based on the order the modes were presented to the listener. This post hoc test was done using a Tukey honestly significant difference (HSD) procedure as all pairwise comparisons were of interest. The second post hoc analysis was done to examine interactions between speaker and presentation mode and was done using *t* tests with a pooled error term. The alpha level was adjusted using a Bonferroni procedure (.05 alpha divided by 8 speakers = .00625).

To measure intralistener reliability, 20% (3 sentences per speaker and mode) were randomly repeated during the transcription of each sentence list in each presentation mode. Sentence transcriptions from both presentations of these sentences were scored independently and then compared to determine consistency of listener performance. A Pearson product–moment correlation coefficient was calculated between scores for the sentences in Presentation 1 versus Presentation 2 ($N = 480$ [3 sentences × 8 speakers × 20 listeners]). Correlations were calculated based on the total number of words correct per sentence in each presentation. Mean difference scores and standard error of measurement were computed for each speaker by determining the differences in scores (based on absolute values) between Presentation 1 and Presentation 2.

Results

Reliability

Pearson product–moment correlation coefficients were .93 for both auditory-only and auditory-visual modes. Mean difference percentage of change (based on absolute values) was 3.16 ($SEM = 1.44$) in the auditory-only mode and 2.78 ($SEM = 1.32$) in the auditory-visual mode.

Intelligibility

Results of the ANOVA are presented in Table 2 for the main effects and two-way interactions. None of the three- or four-way interactions were statistically significant. Significant main effects were found for presentation mode and speaker. The main effects of listener experience and order

TABLE 2. Analysis of variance results across all variables and interactions between variable.

Variables	*df*	*F*	*p*	η^2
Within subjects				
Mode	1	7575.78	<.001*	.670
Speaker	7	147.45	<.001*	.902
Error	126			
Between subjects				
Listener experience	1	1535.6	.133	.136
Order	1	193.7	.581	.019
Error	18			
Interaction				
Speaker × Listener experience	7	1.724	.110	.097
Mode × Listener experience	1	.038	.848	.002
Speaker × Mode	7	4.697	<.001*	.227
Order × Mode	1	7.528	.014*	.320

Note. Asterisk indicates significant results.

were not statistically significant. The significant interactions were between order and presentation mode and speaker and presentation mode.

A post hoc analysis examining differences within and across presentation mode based on listening order revealed that the difference between the two scores of listeners who completed an auditory-only (Session 1) and auditory-visual (Session 2) order was significant (Tukey HSD, $p < .05$). The mean difference was 11.28%, with the auditory-visual scores being higher than the auditory-only scores. Similarly, the difference between the two scores of listeners who completed an auditory-visual (Session 1) and auditory-only (Session 2) order was also significant (Tukey HSD, $p < .05$). The mean difference was 8.17%, with auditory-visual scores being higher than auditory-only scores. The other pairwise comparison that was statistically significant was auditory-only (Session 1) and auditory-visual (Session 1; Tukey HSD, $p < .05$). The mean difference between auditory-only and auditory-visual intelligibility in Session 1 was 14.1%, with auditory-visual being higher. The mean difference between auditory-only intelligibility in Sessions 1 and 2 was 6.23% and was not significant; the mean difference between auditory-visual intelligibility in Sessions 1 and 2 was 3.13% and was not significant. The final comparison of auditory-visual (Session 2) and auditory-only (Session 2) was also not statistically significant ($M = 5.05\%$).

A second post hoc analysis examined the statistically significant interaction between speaker and presentation mode. Post hoc tests revealed a statistically significant difference in intelligibility scores between the two presentation modes (auditory-visual vs. auditory-only) for 3 of the 8 speakers: Speaker A = $t(126) = 4.42$, $p < .001$; Speaker B = $t(126) = 6.68$, $p < .001$; and Speaker C = $t(126) = 4.87$, $p < .001$. The mean difference between presentation modes of these three speakers was 17.6% ($SD = 4.61$). The mean difference between presentation modes for the remaining 5 participants was 4.78% ($SD = 1.31$), with mean auditory-visual scores higher than mean auditory-only scores in all cases, as seen in Figure 1.

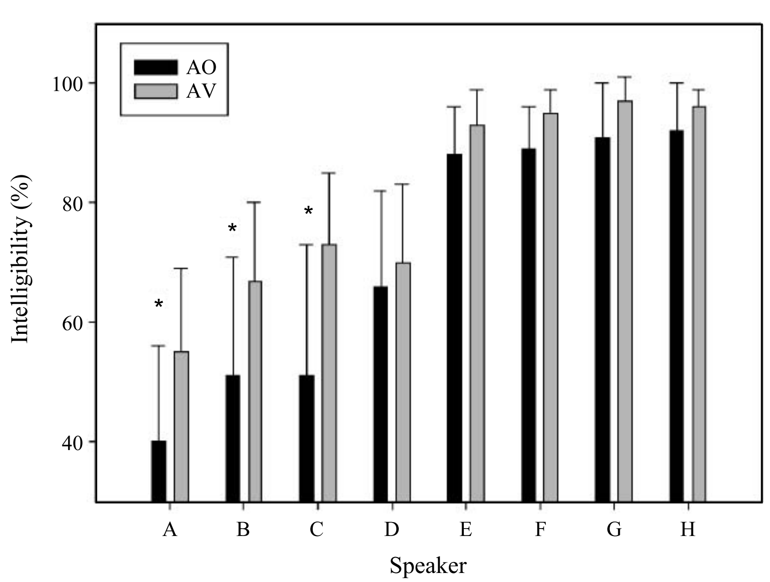
Discussion

The purpose of this study was to examine the influence of visual information on the intelligibility of dysarthric speech in speakers with Parkinson's disease. Although mean intelligibility scores were higher in the auditory-visual mode than in the auditory-only mode for all 8 speakers, these differences were only statistically significant for the 3 speakers with the poorest intelligibility. In addition, familiarity with the speech stimuli had an effect, whereas listener experience did not. These observations can be used to draw clinical implications regarding evaluation and treatment of clients with impaired speech intelligibility.

Influence of Visual Information on Intelligibility

Only those speakers with the lowest intelligibility scores earned significantly higher intelligibility scores when visual information was provided to the listener. This finding is similar to that of Hustad and Cahill (2003), who reported that only the most severely impaired (and least intelligible) of their 5 participants with dysarthria demonstrated a statistically significant increase in intelligibility when visual information was added to the auditory signal. Taken together, results of these two studies suggest that listeners may rely more on visual cues to help them identify a speaker's message when the acoustic signal contains missing, ambiguous, or unclear information. In other words, severity of the intelligibility impairment is an important variable and one that may cause listeners to modify their typical

FIGURE 1. Intelligibility scores for individual speakers for auditory-only (AO) and auditory-visual (AV) modes. Asterisk indicates statistically significant differences.



perceptual strategies. This is consistent with Lindblom's (1990) model of mutuality wherein he proposed that a listener presented with a compromised acoustic signal can understand more of the speaker's message if signal-independent information is available. Nevertheless, it should be noted that Speaker D in the present study represented an exception. Although Speaker D had an intelligibility score (auditory-only) that fell in a moderate to severe range (66%), his score improved only minimally (4%) when visual information was added. Why this speaker showed such a different intelligibility profile from the other speakers is not clear.

Of the 5 speakers who did not demonstrate significant differences in intelligibility scores between the two presentation modes, 4 had intelligibility scores in the auditory-only condition that were above 88%. This suggests that there is a ceiling effect in which the inclusion of visual information may not add relevant cues for listeners. This finding was similar to that of Hustad and Cahill (2003), who observed that their more intelligible speakers' scores did not increase significantly when visual information was added.

Listener Comments on the Inclusion of Visual Information

Most listeners expressed an awareness of an advantage of being able to see the speaker's face during the transcription task. After the two listening sessions were completed, listeners were asked, "What was your experience completing this task over both of these sessions?" Fourteen of the 20 listeners indicated that the task was "easier" or "less frustrating" in the auditory-visual mode, and all but one of these listeners had higher transcription scores in the auditory-visual mode than in the auditory-only mode. Given that half of these 14 listeners had the auditory-only mode first and half had the auditory-visual mode first, the order of presentation did not influence the nature of these postexperimental comments. One listener reported that when she was only listening, she "missed articles, like *the* or *an* at the beginning of sentences," but she said that she was able to detect them when the speaker's face was visible. Another listener stated that "the sound was more like a backup when the visual information was there." Six listeners reported that they focused on the speakers' lips and mouth, and one reported that he benefited from seeing the speaker's expressions.

Four of the 20 listeners reported that the visual image helped with some speakers but not with others. Despite this comment, these listeners all had higher transcription scores for all 8 speakers in the auditory-visual mode than they did in the auditory-only mode. One listener indicated that the auditory-visual condition was "more tiring" and that it was easier to tune into the speaker's voice when not watching the speaker. Contrary to her perception, this listener (like the other listeners) had higher scores when the visual information was provided than when it was not.

Influence of Familiarity on Intelligibility

The statistically significant interaction between presentation mode and order suggests that listener familiarity with the task, the speakers, and/or the stimuli influenced the intelligibility scores. Barkmeier (1988) also reported an interaction between order and presentation mode. In that study, intelligibility scores were significantly higher when visual information was provided, but only if the auditory-visual condition came before the auditory-only condition. A similar (descriptive) observation was made by Hustad and Cahill (2003).

Despite the obvious effect of familiarity on intelligibility, familiarity cannot account for the benefit offered by the addition of visual information in the present study. This can be most clearly seen in the pairwise comparison of the auditory-visual and auditory-only conditions during Session 1, when listeners were unfamiliar with the task, speakers, and stimuli. This comparison indicates that listeners were able to understand speech significantly better when they were able to see and hear the speaker compared to when they could only hear the speaker. The mean intelligibility score across speakers in the auditory-only mode (Week 1) was 67.8%, compared to the mean for speakers in the auditory-visual mode (Week 1) of 82.2%. This was true for both experienced and inexperienced listeners.

Influence of Listener Experience on Intelligibility

There were no significant differences in transcription scores between experienced and inexperienced listeners in this study. This agrees with the findings of Hunter et al. (1991) but disagrees with those of Barkmeier (1988). The reason for these inconsistencies is not clear. Nevertheless, it can be speculated that differences in the speaker group may have influenced listeners' performances. The current study and that of Hunter and colleagues included speakers with a single etiology, whereas the Barkmeier study included speakers with several etiologies. Perhaps it is easier for experienced listeners to switch between speakers with differing perceptual features than it is for inexperienced listeners.

Use of a Dual-Task Paradigm

During pilot work for the present study, speakers with Parkinson's disease were observed to be much more intelligible while being recorded than while they were speaking conversationally prior to the recording session. This was not entirely unexpected in that large differences in intelligibility across speaking conditions (i.e., clinic/laboratory setting vs. spontaneous) have been reported in the literature for individuals with Parkinson's disease (Aronson, 1990; Sarno, 1968; Weismer, 1984). Such improvement in performance is thought to be due to the presence of simple cues that focus a speaker's attention on the target behavior (e.g. recording equipment). There is also evidence in the limb motor control literature that increased attention alone is successful in normalizing many of the deficits of performance associated with Parkinson's disease such as gait and balance (Morris et al., 1996; Oliveira, Gurd, Nixon, Marshall, & Passingham, 1998). The problem of "atypically good intelligibility" was addressed in the present study by enlisting a dual-task paradigm that required the participants to perform manual movements while they were speaking.

When using dual-task paradigms, the underlying assumption is that if attentional capacity is limited and attention is divided between two tasks performed simultaneously, then performance on one or both will be negatively affected (Kahneman, 1973; Wickens, 1984). Dromey and Benson (2003) have suggested that everyday communication may be a type of divided attention. Communication frequently requires speakers to coordinate the demands of message formulation and production with other daily activities such as driving, walking, or watching television. Thus, intelligibility testing as it is usually carried out may create an unrealistic situation by allowing the speaker to focus attention entirely on speech production. Use of a dual-task paradigm during intelligibility testing in this population may actually create a more natural speaking situation and therefore offer a more realistic assessment of a speaker's intelligibility in everyday life contexts.

In a separate study, Bunton and Keintz (2006) reported comparable intelligibility scores during a structured task (sentence reading) and conversational speech in speakers with Parkinson's disease. It is also relevant to note that Bunton and Keintz found no evidence of interference (e.g., entrainment) or disturbances in suprasegmental production or disturbances in language formulation in the dual-task condition compared to a single-task condition (speaking only). Although no measures of manual motor activity were obtained in the present study, the similarity in the tasks, participants, and nature of information observed between the present study and that of Bunton and Keintz strongly suggests that the same was true for the present participants. It is likely that the short duration of the speaking task (single sentences) and the relatively longer duration of pauses between sentences may have prevented entrainment from happening.

Clinical Implications

The present findings have implications for evaluation and management of intelligibility in clients with Parkinson's disease. To begin, this study and others have demonstrated that the provision of visual information is especially beneficial to speakers with severely impaired intelligibility. Nevertheless, the possibility that visual cues could provide clinical benefit for speakers with less severe intelligibility impairments should not be discounted. For example, Speaker G in the present study demonstrated an intelligibility increase of 7% when visual information was added to the auditory information. While the group result wasn't statistically significant, it is possible that 7% improvement could represent a clinically significant improvement in the speaker's functional communication.

Although most evidence suggests that visual information may either benefit or not affect intelligibility (depending on the severity of intelligibility impairment), it is also possible for visual information to hinder intelligibility in speakers with other forms of dysarthria. For example, Meyerson and Fouschee (1978) reported on 22 patients with dysarthria related to Mobius syndrome, a condition commonly associated with bilateral facial paralysis due to involvement of the facial nerve. Some of the participants in this sample were found to use compensatory movements to create near-normal acoustic output for what typically would be bilabial productions. In such cases, the visual information provided by these compensatory movements would not fit with what is being heard. For example, De Feo and Schaefer (1983) reported on a child with Mobius syndrome who produced "bilabial" sounds using a variety of lingual adjustments, including lingual-dental, lingual-buccal, and tip-dental placements. This study and others have shown that listeners identify these compensated bilabial sounds with greater accuracy when visual information is not available than when it is (DeFeo & Schaefer, 1983; Nelson & Hodge, 2000; Von Berg, Brancamp, McColl, & Von Berg, 2005). Thus, in speakers with Mobius syndrome, it appears that access to visual information distracts listeners from the (relatively accurate) acoustic signal and may thereby have a negative effect on their speech intelligibility.

As suggested by Kent, Miolo, and Bloedel (1994), each speaker has a range of potential intelligibility that is influenced by factors related to the speaker, listener, and conditions in which communication takes place. Given that visual cues are commonly available to listeners in most everyday communication situations, it may be appropriate to include visual information when evaluating a client's functional intelligibility. By evaluating intelligibility both with and without visual information, it may be possible to determine the best management strategy for individual clients. For example, if visual information benefits intelligibility for a given client, management might focus on helping the client and his or her communication partner maximize the use of visual cues. Findings of the studies examining the influence of visual information provide clinicians with another source of information that may benefit clients with impaired intelligibility.

Acknowledgments

The authors wish to thank Dr. Julie Barkmeier-Kraemer and Dr. Brad Story for their valuable input to the design of this study and for their suggestions on an earlier version of this article. Support for this research was provided by National Multipurpose Research and Training Center Grant DC-01409 from the National Institute on Deafness and Other Communication Disorders (NIDCD), NIDCD/National Institutes of Health Grant RO3-DC005902, and the Douglas G. Stuart Predoctoral Fellowship in Neuroscience at the University of Arizona. We thank Becca Lambson and Joanne McIntyre for video editing and Dr. Mark Borgstrom for statistical support. We are grateful to the speakers who produced the stimuli and the listeners who participated in the experiments.

References

Adobe Systems. (2004). Adobe Premiere Pro [Computer software]. Seattle, WA: Author.

American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening*. Available from www.asha.org/policy.

Ansel, B. A. (1985). *Acoustic predicators of speech intelligibility in cerebral palsied dysarthric adults*. Unpublished doctoral dissertation, University of Wisconsin—Madison.

Aronson, A. E. (1990). *Clinical voice disorders* (3rd ed.). New York: Thieme-Stratton.
Barkmeier, J. M. (1988). *Intelligibility of dysarthric speakers: Audio-only and audio-visual presentations*. Unpublished master's thesis, University of Iowa, Iowa City.
Bench, J., & Bamford, J. (1979). *Speech-hearing tests and the spoken language of hearing-impaired children*. London: Academic Press.
Berry, R. A., & Knight, R. E. (1975). Auditory versus audio-visual intelligibility measurements of alaryngeal speech: A preliminary report. *Perceptual and Motor Skills, 40,* 915–918.
Beukelman, D. R., & Yorkston, K. M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. *Journal of Communication Disorders, 13,* 33–41.
Bunton, K., & Keintz, C. K. (2006, March). *Effects of a concurrent motor task on speech intelligibility for speakers with Parkinson disease*. Paper presented at the Conference on Motor Speech, Austin, TX.
Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research, 12,* 462–496.
Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12,* 246–269.
De Feo, A. B., & Schaefer, C. M. (1983). Bilateral facial paralysis in a preschool child: Oral-facial and articulatory characteristics (a case study). In W. R. Berry (Ed.), *Clinical dysarthria* (pp. 165–186). San Diego, CA: College Hill Press.
Dromey, C., & Benson, A. (2003). Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research, 46,* 1234–1246.
Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management* (2nd ed.). St. Louis, MO: Elsevier Mosby.
Duffy, J., & Giolas, T. (1974). Sentence intelligibility as a function of key word selection. *Journal of Speech and Hearing Research, 17,* 631–637.
Enderby, P. M. (1983). *The Frenchay Dysarthria Assessment*. San Diego, CA: College Hill Press.
Garcia, J. M., & Cannito, M. P. (1996). Influence of verbal and nonverbal context on the sentence intelligibility of a speaker with dysarthria. *Journal of Speech and Hearing Research, 39,* 750–760.
Ho, A. K., Iansek, R., & Bradshaw, J. L. (2002). The effect of a concurrent task on Parkinsonian speech. *Journal of Clinical and Experimental Neuropsychology, 24*(1), 36–47.
Hubbard, D. J., & Kushner, D. (1980). A comparison of speech intelligibility between esophageal and normal speakers via three modes of presentation. *Journal of Speech and Hearing Research, 23,* 909–916.
Hunter, L., Pring, T., & Martin, S. (1991). The use of strategies to increase speech intelligibility in cerebral palsy: An experimental evaluation. *Journal of Speech and Hearing Research, 26,* 163–174.
Hustad, K. C., & Cahill, M. A. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology, 12,* 198–208.
Hustad, K. C., Jones, T., & Dailey, S. (2003). Implementing speech supplementation strategies: Effects of intelligibility and speech rate of individuals with chronic severe dysarthria. *Journal of Speech, Language, and Hearing Research, 46,* 462–474.
Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
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,* 1339–1351.
Kent, R. D., Miolo, G., & Bloedel, S. (1994). The intelligibility of children's speech: A review of evaluation procedures. *American Journal of Speech-Language Pathology, 3,* 81–95.
Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders, 54,* 482–499.
Lidestam, B., & Beskow, J. (2006). Visual phonemic ambiguity and speechreading. *Journal of Speech, Language, and Hearing Research, 49,* 835–847.
Lindblom, B. (1990). On the communication process: Speaker-listener interaction and the development of speech. *Augmentative and Alternative Communication, 6,* 220–230.
MacLeod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and auditory-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology, 24,* 29–43.
Massaro, D. W. (1987). *Speech perception by eye and ear: A paradigm for psychological inquiry.* Hillsdale, NJ: Erlbaum.
Menke, E. O., Oschner, G. J., & Testut, E. W. (1983). Listener judges and the speech intelligibility of deaf children. *Journal of Communication Disorders, 16,* 175–180.
Meyerson, M. D., & Fouschee, D. R. (1978). Speech, language, and hearing in Moebius syndrome: A study of 22 patients. *Developmental Medicine and Child Neurology, 20,* 357–365.
Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders, 48,* 286–296.
Morrell, C. H., Gordon-Salant, S., Pearson, J. D., Brant, L. J., & Fozard, J. L. (1996). Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *Journal of the Acoustical Society of America, 100,* 1949–1967.
Morris, M., Iansek, R., Matyas, T., & Summers, J. (1996). Stride length regulation in Parkinson's disease: Normalising strategies and underlying mechanisms. *Brain, 119,* 551–568.
Neely, K. K. (1956). Effects of visual factors on the intelligibility of speech. *Journal of the Acoustical Society of America, 28,* 1275–1277.
Nelson, M. A., & Hodge, M. M. (2000). Effects of facial paralysis and audiovisual information on stop place identification. *Journal of Speech, Language, and Hearing Research, 43,* 158–171.
Oliveira, R., Gurd, J., Nixon, P., Marshall, J., & Passingham, R. (1998). Hypometria in Parkinson's disease: Automatic versus controlled processing. *Movement Disorders, 13,* 422–427.
O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech and Hearing Disorders, 19,* 429–439.
Platt, L. J., Andrews, G., Young, M., & Quinn, P. T. (1980). Dysarthria of adult cerebral palsy: I. Intelligibility and articulatory impairment. *Journal of Speech and Hearing Research, 23,* 28–40.
Rosen, S. M., & Corcoran, T. (1982). A video-recorded test of lip-reading for British English. *British Journal of Audiology 16,* 245–254.
Rosenblum, L. D., Johnson, J. A., & Saldana, H. M. (1996). Point-light facial displays enhance comprehension of speech in noise. *Journal of Speech and Hearing Research, 39,* 1159–1170.
Sarno, M. (1968). Speech impairment in Parkinson's disease. *Archives of Physiological Medicine and Rehabilitation, 49,* 269–275.

Siegenthaler, B. M., & Gruber, V. (1969). Combining vision and audition for speech reception. *Journal of Speech and Hearing Disorders, 34,* 58–60.
Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26,* 212–215.
Summerfield, A. Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 3–51). Hove, United Kingdom: Erlbaum.
Tikofsky, R. S., & Tikofsky, R. P. (1964). Intelligibility measures of dysarthric speech. *Journal of Speech and Hearing Research, 7,* 325–333.
Von Berg, S., Brancamp, T., McColl, D., & Von Berg, S. (2005, November). *Moebius syndrome: Measures of listener comprehensibility versus intelligibility.* Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, San Diego, CA.
Weismer, G. (1984). Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. In M. McNeil, J. Rosenbek, & A. Aronson (Eds.), *The dysarthrias: Psychology, acoustics, perception, management* (pp. 101–130). San Diego, CA: College Hill Press.
Wickens, C. (1984). Processing resources in attention. In R. Parasuraman & D. Davies (Eds.), *Varieties of attention* (pp. 63–101). Orlando, FL: Academic Press.
Yorkston, K. M., & Beukelman, D. R. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders, 11,* 499–512.
Yorkston, K. M., & Beukelman, D. R. (1980). A clinician-judged technique for quantifying dysarthric speech based on single-word intelligibility. *Journal of Communication Disorders, 13,* 15–31.
Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of intelligibility of dysarthric speech*. Tygard, OR: C.C. Publications.
Yorkston, K. M., Beukelman, D. R., & Tice, R. (1996). Sentence Intelligibility Test for Macintosh [Computer software]. Lincoln, NE: Tice Technology Services.

Received October 4, 2005
Revision received March 6, 2006
Accepted January 30, 2007
DOI: 10.1044/1058-0360(2007/027)

Contact author: Connie Keintz, who is now at Florida Atlantic University, Department of Communication Sciences and Disorders, 777 Glades Road, P.O. Box 3091, Boca Raton, FL 33431. E-mail: ckeintz@fau.edu.

Appendix (p. 1 of 2)

Institute of Hearing Research Sentence Stimuli for Speaker Recording Task (MacLeod & Summerfield, 1990)

List 2

1. He tore his shirt.
2. They finished the jigsaw.
3. She brought her camera.
4. The lady watered her plants.
5. The salt cellar's full.
6. The boy hit his thumb.
7. The mother shook her head.
8. The snow lay on the hills.
9. The father used a towel.
10. The tree was in the back garden.
11. The yacht sailed past.
12. The lady pushed the pram.
13. They're leaving today.
14. The picture hung on the wall.
15. The children sit under the tree.

List 3

1. The lunch was very early.
2. The dirty boy is washing.
3. He hid his money.
4. The curtains were too short.
5. The knife cut the cake.
6. They emptied their pockets.
7. The new shoes were tight.
8. The coat hangs in the cupboard.
9. The sun shone through the clouds.
10. She took her purse.
11. The team lost the match.
12. The shirt caught on a nail.
13. They picked some raspberries.
14. The man climbed the mountain.
15. The lady hurt her arm.

List 4

1. The old clothes were dirty.
2. He carried a stick.
3. She read her book.
4. The new house was empty.
5. The thief brought a ladder.
6. The horse stands by the gate.
7. They're heading for the park.
8. The gardener trimmed the hedge
9. They're standing up.
10. Someone's hiding in the bushes.
11. The waiter lit the candles.
12. The baker iced the cake.
13. The woman slipped on the ice.
14. The small puppy was scared.
15. The lady changed her mind.

List 5

1. The daughter closed the box.
2. They broke into the safe.
3. The doctor carries a bag.
4. The new game was silly.
5. The little boy was tired.
6. They saw the sign.
7. She's wrapping the parcel.
8. The children laughed at the clown.
9. The apple pie was hot.
10. The ship sailed up the river.
11. The house has a lovely garden.
12. The noisy dog is barking.
13. They bought some tickets.
14. The man goes to the bank.
15. The nurse helped the child.

List 6

1. The girl knew the story.
2. He reached for a cup.
3. The lady was quite cross.
4. The rope was too short.
5. She's listening to the radio.
6. The husband cleaned the car.
7. The postman leaned on the fence.
8. The china vase was broken.
9. The other team won.
10. They locked the safe.
11. The leaves dropped from the trees.
12. The men watched the race.
13. The bird's building a nest.
14. The woman called her dog.
15. They're waving at the train.

List 7

1. The cat scratched the chair.
2. She tapped at the window.
3. The man painted the gate.
4. He slid on the floor.
5. They're lifting the box.
6. The woman listened to her friend.
7. The driver hooted his horn.
8. The cake tasted nice.
9. The sailor stood on the deck.
10. The young girls were pretty.
11. They painted the ceiling.
12. The back door was shut.
13. The tree lost its leaves.
14. The boy eats with his fork.
15. The young mother's shopping.

Appendix (p. 2 of 2)

Institute of Hearing Research Sentence Stimuli for Speaker Recording Task (MacLeod & Summerfield, 1990)

List 8	List 10
1. The girl sharpened her pencil.	1. She sings in the bath.
2. She closed her eyes.	2. The meat was too tough.
3. The plant grows on the wall.	3. The child ate some jam.
4. The puppy licked his master.	4. They're stealing the apples.
5. The family's having a picnic.	5. The children dried the dishes.
6. The train arrived on time.	6. The paper boy was cheeky.
7. They won the game.	7. The little car was slow.
8. The lady waited for her husband.	8. The bath tubs are dripping.*
9. The post office was near.	9. They came at Easter.
10. They rowed the boat.	10. He's wearing a tie.
11. The old fox was sly.	11. The new towel was clean.
12. The baby lost his rattle.	12. The water poured from a jug.
13. He dug with his spade.	13. The red apples were in a bowl.
14. The boiled egg was soft.	14. The bus stopped at the shops.
15. The two ladies were watching.	15. The man drew with a pencil.

*The word "tubs" was substituted for "taps," which was felt to represent nonstandard English word use while preserving the visual contribution of sounds.