Effect of Fingerspelling Task on Temporal Characteristics and Perceived Naturalness of Speech Produced During Simultaneous Communication by Inexperienced Signers

Robert L. Whitehead
National Technical Institute for the Deaf, Rochester, NY

Douglas MacKenzie
Ashley Broska
Jocelyn Proietti
Emily Vogel
Nicole Williams
Dale E. Metz
Nicholas Schiavetti
State University of New York, Geneseo

Simultaneous communication (SC) that combines speech with manual signs and fingerspelling to produce each word of an utterance is often used to improve the communication of children who are deaf and/or hard of hearing and enhance their speech and language development (Vernon & Andrews, 1990). Vernon and Andrews (1990) stated that the advantages of SC include more accurate representation of English demonstrated increased interword interval, diphthong, word, and sentence durations. Regression analysis indicated significant correlations between temporal measures and rated speech naturalness, and an analysis of variance showed significant increases in segmental and interword interval durations and significant decreases in rated speech naturalness as fingerspelling task length increased. Meta-analysis revealed that during SC, inexperienced signers demonstrated greater temporal elongation than experienced signers, especially in anticipation of fingerspelled words, and greater reductions in speech naturalness.

KEY WORDS: simultaneous communication, fingerspelling, speech naturalness
than that provided by lipreading alone and increased psychosocial and linguistic development. Disadvantages of SC include alterations in the linguistic integrity of both manual and oral forms of communication, abbreviations of English in the manual code, deletion of grammatical markers in signs, and slowing of speech. Although research on SC (Huntington & Watton, 1984; Marmor & Petitto, 1979; Strong & Charlson, 1987) has provided useful information regarding its relative advantages and disadvantages, more research is needed to examine the quality of the speech model presented to deaf and hard-of-hearing children during SC, especially among the growing body of inexperienced signers that children are exposed to at home and in school (Watson, Hardie, Archbold, & Wheeler, 2008; Woodward & Allen, 1993).

Previous research (Huntington & Watton, 1984; Whitehead, Schiavetti, Whitehead, & Metz, 1995; Windsor & Fristoe, 1989, 1991) has indicated that speech produced during SC differs from speech produced alone (SA) in certain temporal features, including a slower rate of articulation and increases in sentence, word, vowel, and interword interval durations. Windsor and Fristoe (1989, 1991) concluded that speech movements become slower during SC in order to maintain simultaneity between speech and manual communication. Whitehead et al. (1995) demonstrated that the addition of fingerspelling to the manual task further elongated speech segment and pause durations during SC. They found that the addition of even a relatively simple fingerspelling task in SC further increased segmental and suprasegmental durations and suggested that future research should consider the effect of such manual variables as number of gestures, sign complexity, and fingerspelling task length on the temporal and perceptual characteristics of speech during SC.

During SC, the communicator typically uses signs the majority of the time but uses fingerspelling, which is slower than signing, for words for which there is no sign. Such fingerspelled words are often orthographically multisyllabic or are composed of many letters and therefore require a more difficult manual task than signing. Evidence of this difficulty has been provided by Akamatsu and Stewart (1989, p. 370), who stated that “hearing people learning to sign report that it is very difficult to learn fingerspelling,” and by Wilcox (1992), who reported that hearing adults who were enrolled in a sign language course rated learning to read fingerspelling as their most difficult learning task. Akamatsu and Stewart also reported that it is difficult for both children and adults to fingerspell and speak simultaneously due to the demanding task of coordinating manual orthographic patterns with articulatory speech patterns. They also noted that the use of fingerspelling in SC has not been sufficiently studied despite its central importance to the development of language and literacy. Therefore, further investigation of the effects of variations in fingerspelling task difficulty on speech characteristics during SC is warranted to provide better generalization to the demands of real-world simultaneous oral and manual communication, especially among people who are just learning manual communication.

One study has examined the systematic effects of fingerspelling task difficulty on the SC performance of experienced signers. Schiavetti, Whitehead, Whitehead, and Metz (1998) studied the effects of SA versus SC and fingerspelling difficulty (indexed as four graduated levels of fingerspelling task length) on temporal variables such as sentence and word duration and examined the influence of temporal increases on rated speech naturalness. Groups of four experimental words were composed to sample increasing fingerspelling task length by beginning with an initial base word and adding three suffixes of graduated increasing length. Results showed longer durations in SC than SA, and sentence and word durations increased systematically with increased fingerspelling task length. Sentence and word duration increases showed steeper sloped functions for SC than for SA, revealing a significant interaction effect between experimental condition and fingerspelling task length, in which the influence of orthographic word length increased the segmental duration more rapidly in SC than in SA. In addition, Schiavetti et al. found that the SC speech was rated more unnatural than SA for all fingerspelling task length levels, and speech naturalness decreased in SC as fingerspelling task length increased, but not in SA. Multiple regression analysis using temporal characteristics of speech as predictor variables and the scaled speech naturalness as the predicted variable demonstrated a highly significant relationship of selected temporal speech production characteristics and speech naturalness, with sentence duration (reflecting overall speech rate) as the strongest predictor variable. The conclusion was quite clear that speech during SC was perceived as significantly more unnatural than SA, and perceived naturalness decreased systematically with fingerspelling task length in the SC mode.

However, no empirical research has been conducted to measure disruption in speech production characteristics or natural speech quality during SC by the inexperienced signers who find fingerspelling so difficult to learn (Akamatsu & Stewart, 1989; Wilcox, 1992), nor has there been any empirical documentation of the relationship between specific temporal speech characteristics and perception of speech quality during SC by inexperienced signers. Because external validity research has not been reported on the effects of fingerspelling task difficulty, systematic replication is needed to determine the generalizability of findings (Pedhazur & Schmelkin, 1991) with experienced signers across persons with less signing experience who are more representative of newly signing parents or teachers in the language learning environments of children who are deaf or hard of hearing. Generalizing results beyond highly experienced signers is important because inexperienced signers such as parents, siblings, peers, or school teachers who use SC with children who are deaf or hard of hearing may demonstrate more speech disruption during SC than experienced signers. In addition, Lodge-Miller and Elfenbein (1994) found that inexperienced signers tend to overestimate their signing ability, implying that they may also overestimate their SC ability and so may need more practice to achieve fluency.

Therefore, the purposes of this study were to examine the effect of SC and the effect of increasing fingerspelling difficulty level on selected temporal characteristics and scaled speech naturalness for inexperienced signers. Specifically,
we studied the effects of two independent variables: (a) mode of communication (SA vs. SC) and (b) fingerspelling difficulty (indexed as four graduated levels of fingerspelling task length from low to high) on six dependent variables: (a) sentence duration, (b) experimental word duration, (c) diphthong duration in the word preceding the experimental word, (d) interword interval duration preceding the experimental word, (e) interword interval duration following the experimental word, and (f) rated speech naturalness. In addition, the correlations between the five temporal characteristics of speech production (dependent variables a through e) and rated speech naturalness (dependent variable f) were examined in a multiple regression analysis that was designed to predict rated speech naturalness from the temporal characteristics.

METHOD

Speakers

Speakers for this study were 8 hearing adult females who were students in the Department of Communicative Disorders and Sciences at the State University of New York, Geneseo. All of the speakers had just successfully completed a one-semester beginning course in sign language, which included the use of speech combined with signed English and fingerspelling, when they were recorded for this study.

Speech Stimuli

The speech samples investigated consisted of the carrier sentence, “I can say _________ again” and 16 fingerspelled experimental words embedded in the blank slot in the carrier sentence. The carrier sentence was constructed to carry a list of experimental words in a phonetic context designed to precede each experimental word with a diphthong and follow each experimental word with a central vowel to aid in locating the experimental words in a digitized acoustic recording for temporal analysis.

Four groups of four experimental words each, with each group consisting of words of increasing fingerspelling task length, were used as an index of fingerspelling difficulty. An increase in fingerspelling task length was operationally defined as an increase in the number of letters orthographically displayed per word. Each group began with an initial base word to which three suffixes of increasing length were added so as to increase the base word’s length. Thus, Level 1 of task length consisted of base forms of the four experimental words, and Levels 2, 3, and 4 of task length consisted of the base forms with suffixes of increasing length added. The stimulus material consisted of the four base words, care, talk, truth, and trust and their expansions into words of increasing length by the addition of suffixes (e.g., care, careless, carelessly, carelessness). This task is similar to the one employed by Lehiste (1972) in a study of utterance timing in which she found that the addition of longer suffixes increased whole-word utterance length while decreasing the length of the base word within the whole word. Thus, it was expected that word and sentence duration would increase with suffix length in both the SA and SC conditions and that a greater relative increase in duration for SC would reflect the effect of fingerspelling task length on temporal speech characteristics. Table 1 presents the four groups of four experimental words each.

Recording Procedure

The 8 speakers were shown a flash card of each sentence and were asked to read the sentence. Audio recordings were made in a sound-treated booth using an Audio-Technica AT-816 microphone that was placed 15 cm from each speaker’s mouth. The microphone was connected to a Tascam 202MKII tape deck. The speakers produced each group of sentences under two conditions: (a) SA and (b) SC using speech combined with signed English and fingerspelling. The sentences containing the 16 experimental words were presented to the speakers on flashcards in two different random orders: one for SA and one for SC. The order of experimental condition (SA vs. SC) was counterbalanced across speakers. The speakers were shown the experimental words before the recording so they could familiarize themselves with the signs and fingerspellings to be used in SC. During data collection, each speaker was carefully monitored to ensure that she produced the signs and fingerspellings correctly.

Acoustic Analysis Procedure

For each speech sample, duration measures in milliseconds were determined for (a) the entire sentence, (b) the experimental word, (c) the diphthong within the word say preceding the experimental word in the carrier sentence, (d) the interword interval preceding the experimental word, and (e) the interword interval following the experimental word. All duration measures were made by a trained research assistant under the direct supervision of one of the authors (DEM).

The acoustic signal from the audio recording of each sentence was digitized with 16-bit precision at a sampling rate of 20 kHz using Kay Elemetrics Computerized Speech Lab (CSL Model 4300B). When the digitizing process is initiated, the Computerized Speech Lab applies an appropriate internal low-pass anti-aliasing filter to the raw acoustic signal (at a sampling rate of 20 kHz, the upper frequency cutoff is 8 kHz), stores the digitized results in memory, and displays the resulting waveform on a graphics monitor.

| Experimental words organized into groups according to fingerspelling task length. |
|-------------------------------|----------------|----------------|----------------|----------------|
|                             | 1              | 2              | 3              | 4              |
| care                          | careless       | carelessly     | carelessness   |
| talk                          | talkative      | talkatively    | talkativeness  |
| truth                         | truthful       | truthfully     | truthfulness   |
| trust                         | trustful       | trustfully     | trustfulness   |
For each experimental sentence and for each of the five duration measures, cursors on the Computerized Speech Lab were placed in the appropriate location on the acoustic waveform, and the duration, in milliseconds, was obtained.

**Sentence duration.** For measurement purposes, sentence duration was defined as the time from the initiation of phonation evidenced by the initiation of increased amplitude and periodicity of the acoustic waveform for the diphthong /ai/ in the word I to the termination of phonation and decreased signal amplitude for the consonant /n/ in the word again.

**Word duration.** Word duration was defined as the time from the initiation of initial plosive phoneme to the termination of final phoneme in each experimental word. This was identified by first locating the initial plosive burst for each experimental word. The termination of the final phoneme for each experimental word was determined by identifying one of the following: (a) plosive burst for words ending in /l/, /r/, or /k/; (b) termination of signal amplitude and periodicity for words ending in /v/, /i/, or /l/; (c) termination of voicing (signal periodicity) for words ending in /v/; or (d) termination of noise (signal aperiodicity) for words ending in /s/ or /l/.

**Diphthong duration.** Diphthong duration in the word say was identified as the time from the point where the acoustic signal changed from the noise (aperiodicity) for the /s/ to the increased amplitude and signal periodicity for the diphthong /ei/ to the point where signal amplitude and periodicity ended.

**Interword intervals.** Interword interval before the experimental word was determined by placing one cursor on the termination of the diphthong /ei/ in say and the other on the plosive burst of the initial phoneme of the experimental word. Interword interval after the experimental word was determined by placing one cursor on the terminal end of the experimental word and the other cursor on the initiation of phonation /s/ in again as identified by increased amplitude and periodicity of the acoustic signal.

**Reliability of Acoustic Analysis**

**Intra-observer reliability.** As a measure of intra-observer reliability of the acoustical measurement procedure, the recordings of two speakers were selected at random, and all five acoustic temporal characteristics described above were measured a second time by the same person who made the first measurement. The correlations between the original and remeasured sets of SA data were 0.99 for sentence duration, 0.98 for word duration, 0.96 for diphthong duration, 0.97 for interword interval before duration, and 0.98 for interword interval after duration. These high correlations indicate substantial intra- and interobserver reliability for the acoustic duration measures.

**Scaling of Perceived Speech Naturalness**

The perceived naturalness of the speech samples was scaled by listeners using the 9-point equal-appearing interval scale developed by Martin, Haroldson, and Triden (1984).

**Naturalness listeners.** Forty communicative disorders students at State University of New York, Geneseo served as listeners to scale the speech naturalness of the recordings. All listeners passed a hearing screening at 20 dB HL (American National Standards Institute, 2004) at .5, 1, and 2 kHz; spoke English as their first language; and were unaware of the experimental variables in the speech recordings they scaled.

**Scaling materials and procedures.** The 40 listeners were randomly assigned into groups of five to audit the recordings of each of the 8 speakers. Each listener group scaled the naturalness of 32 sentences (4 word groups × 4 levels of fingerspelling task length × 2 experimental conditions) arranged in a different random order for each speaker. The scaling procedures and instructions were the same as those employed by Martin et al. (1984). Recordings of the speakers were played to the listeners, who were seated in a sound-attenuating room in seats arranged along an arc traced at 2 m from the center of the playback speaker. Listeners were given the instructions outlined by Martin et al., in which a definition of naturalness is not imposed on the listeners by the experimenter and where 1 corresponds to highly natural-sounding speech and 9 corresponds to highly unnatural-sounding speech. The interval scaling task was illustrated to the listeners by showing them a picture of several parallel horizontal lines ranging in length from 0.32 cm to 25 cm and demonstrating how the perceived length of the lines could be scaled with an interval scaling procedure. The listeners practiced the scaling procedure by scaling the length of several lines, as suggested by Stevens (1975). The listeners then listened to four examples (two SA and two SC) to familiarize them with the range of speech rates they would audit, but they were not told what speech characteristics to attend to in the sample.

**Analysis of naturalness scaling data.** The arithmetic means of the listeners’ interval scale data were computed for each of the sentences for each speaker. Means and standard deviations were then calculated for these group mean data for the SA and SC speech samples at each of the four levels of fingerspelling complexity.

**Reliability of naturalness measurements.** To avoid the error associated with listener memory, the reliability of the
interval scale naturalness rating measurements was analyzed with the intraclass correlation (Ebel, 1951) instead of with test–retest correlations as suggested by Guilford (1954, p. 395). The reliability coefficients for the group and individual scaling data were 0.96 and 0.81, respectively. These reliability coefficients are consistent with speech naturalness data from a number of studies with different populations (Martin et al., 1984; Metz, Schiavetti, & Sacco, 1990; Schiavetti, Martin, Haroldson, & Metz, 1994; Schiavetti, Whitehead, Whitehead, & Metz, 1998). The group reliability coefficient was very high, and the individual reliability coefficient was somewhat lower, as is the usual case with intraclass correlations for such scaling data (Guilford, 1954). These intraclass correlations indicate that the single listener reliability would be weaker than the group reliability if used to rate speech naturalness. Because group mean data such as these naturalness ratings are commonly employed for research, the group reliability coefficients more accurately reflect the reliability of data used to determine speech naturalness in this and the other reported studies.

Statistical Analysis of Acoustic and Perceptual Data

Statistical analysis of the acoustic and perceptual data was accomplished by a 2 × 4 analysis of variance (ANOVA) with repeated measures on both independent variables: (a) experimental condition (SA vs. SC) and (b) fingerspelling task length (four levels numbered 1 to 4). One ANOVA was employed for each of the six dependent variables: (a) sentence duration, (b) word duration, (c) diphthong duration in the word say before the experimental word, (d) interword interval duration before the experimental word, (e) interword interval duration after the experimental word, and (f) scaled speech naturalness. The cell entry datum for each dependent variable was the mean across the four target word roots (i.e., care, talk, truth, and trust) for each speaker in each communication mode and task length condition. The degrees of freedom in the ANOVAAs, then, were based on a data matrix with the intraclass correlation (Ebel, 1951) instead of with test–retest correlations as suggested by Guilford (1954, p. 395). The reliability coefficients for the group and individual scaling data were 0.96 and 0.81, respectively. These reliability coefficients are consistent with speech naturalness data from a number of studies with different populations (Martin et al., 1984; Metz, Schiavetti, & Sacco, 1990; Schiavetti, Martin, Haroldson, & Metz, 1994; Schiavetti, Whitehead, Whitehead, & Metz, 1998). The group reliability coefficient was very high, and the individual reliability coefficient was somewhat lower, as is the usual case with intraclass correlations for such scaling data (Guilford, 1954). These intraclass correlations indicate that the single listener reliability would be weaker than the group reliability if used to rate speech naturalness. Because group mean data such as these naturalness ratings are commonly employed for research, the group reliability coefficients more accurately reflect the reliability of data used to determine speech naturalness in this and the other reported studies.

Table 2. Means and standard deviations (SDs) of the five temporal variables and speech naturalness during speech alone and simultaneous communication conditions at four different levels of fingerspelling task length.

<table>
<thead>
<tr>
<th>Condition and fingerspelling task length</th>
<th>Sentence duration</th>
<th>Word duration</th>
<th>Diphthong duration</th>
<th>IWI before duration</th>
<th>IWI after duration</th>
<th>Speech naturalness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Speech alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1509</td>
<td>134</td>
<td>417</td>
<td>81</td>
<td>153</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>1720</td>
<td>141</td>
<td>582</td>
<td>34</td>
<td>160</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td>1849</td>
<td>175</td>
<td>721</td>
<td>53</td>
<td>169</td>
<td>61</td>
</tr>
<tr>
<td>4</td>
<td>1894</td>
<td>120</td>
<td>761</td>
<td>57</td>
<td>173</td>
<td>52</td>
</tr>
<tr>
<td>Simultaneous communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4252</td>
<td>890</td>
<td>633</td>
<td>112</td>
<td>360</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>6093</td>
<td>1240</td>
<td>2018</td>
<td>967</td>
<td>357</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>6890</td>
<td>1215</td>
<td>2773</td>
<td>1532</td>
<td>345</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>8246</td>
<td>1918</td>
<td>3439</td>
<td>1490</td>
<td>351</td>
<td>52</td>
</tr>
</tbody>
</table>

Note. All temporal durations are in milliseconds (IWI = interword interval). Speech naturalness is on a 1–9-point interval scale, with 1 = most natural and 9 = most unnatural.

RESULTS

Descriptive Statistics

The results of this experiment are presented in Table 2, which shows the means and standard deviations of the dependent variables in each of the experimental conditions, and in Figures 1 through 6, which demonstrate the effects of the independent variables on each of the dependent variables.

Effect of Communication Mode and Fingerspelling Task Length on Acoustic Variables

Figure 1 reveals that SC was longer in sentence duration than SA for all task length levels, F(1, 7) = 114.93, p < 0.0001, η² = 0.94, and that sentence duration increased systematically with increased fingerspelling task length, F(3, 21) = 68.55, p < 0.0001, η² = 0.90. The steeper slope of the function for SC than for SA shown in Figure 1 reveals a significant interaction effect between experimental condition and fingerspelling task length, F(3, 21) = 48.42, p < 0.0001, η² = 0.87. Bonferroni t tests revealed significant differences between Levels 1 and 2 and between Levels 2 and 3 for both SA and SC, but no significant differences between Levels 3 and 4 for either condition.

Figure 2 reveals that SC was longer in word duration than SA for all task length levels, F(1, 7) = 21.87, p < 0.01, η² = 0.75, and that word duration increased systematically with increased fingerspelling task length, F(3, 21) = 23.24, p < 0.01, η² = 0.79. The steeper slope of the
function for SC than for SA shown in Figure 2 reveals a significant interaction effect between experimental condition and fingerspelling task length, $F(3, 21) = 16.72$, $p < 0.01$, $\eta^2 = 0.70$. Bonferroni $t$ tests revealed significant differences between Levels 1 and 4 and between Levels 2 and 4 of fingerspelling task length for SC but no significant differences among the levels for SA.

Figure 3 reveals that SC was longer in diphthong duration than SA for all task length levels, $F(1, 7) = 275.74$, $p < 0.0001$, $\eta^2 = 0.97$, that diphthong duration did not appear to vary systematically with fingerspelling task length, $F(3, 21) = 0.18$, $p = 0.90$, $\eta^2 = 0.02$, and that there was no apparent interaction effect, $F(3, 21) = 3.5$, $p = 0.03$, $\eta^2 = 0.33$.

Figure 4 reveals that SC was longer in the duration of the interword interval before the experimental word than SA for all task length levels, $F(1, 7) = 57.92$, $p < 0.0001$, $\eta^2 = 0.89$, and that interword interval duration before the experimental word increased systematically with increased fingerspelling task length, $F(3, 21) = 8.9$, $p < 0.001$, $\eta^2 = 0.56$. The slightly steeper slope of the function for SC than for SA shown in Figure 4, especially from Level 3 to Level 4, reveals a significant interaction effect between experimental condition and fingerspelling task length, $F(3, 21) = 9.91$, $p < 0.001$, $\eta^2 = 0.58$. Bonferroni $t$ tests revealed significant differences between Levels 1 and 4 and between Levels 2 and 4 of fingerspelling task length for SC but no significant differences among the levels for SA.

Figure 5 reveals that SC was longer in the duration of the interword interval after the experimental word than SA for all task length levels, $F(1, 7) = 15.83$, $p < 0.01$, $\eta^2 = 0.69$, that interword interval after the experimental word did not appear to vary systematically with fingerspelling task length, $F(3, 21) = 1.64$, $p = 0.21$, $\eta^2 = 0.19$, and that there was no apparent interaction effect, $F(3, 21) = 1.60$, $p = 0.22$, $\eta^2 = 0.18$.

Effect of Communication Mode and Fingerspelling Task Length on Speech Naturalness

Figure 6 reveals that SC was rated more unnatural on the speech naturalness scale than SA for all fingerspelling task length levels, $F(1, 7) = 136.90$, $p < 0.0001$, $\eta^2 = 0.95$, and that speech became more unnatural as fingerspelling task length increased, $F(3, 21) = 15.60$, $p < 0.0001$, $\eta^2 = 0.69$. This effect of task length, however, was only applicable for the SC condition, as indicated by a significant interaction
Figure 3. Mean diphthong duration and standard deviation in milliseconds plotted against fingerspelling task length level with communication mode (speech alone vs. simultaneous communication) as the parameter.

Figure 4. Mean interword interval before experimental word duration and standard deviation in milliseconds plotted against fingerspelling task length level with communication mode (speech alone vs. simultaneous communication) as the parameter.

Figure 5. Mean interword interval after experimental word duration and standard deviation in milliseconds plotted against fingerspelling task length level with communication mode (speech alone vs. simultaneous communication) as the parameter.

Figure 6. Mean interval scale value and standard deviation of rated speech naturalness (1 = natural and 9 = unnatural) plotted against fingerspelling task length level with communication mode (speech alone vs. simultaneous communication) as the parameter.
between experimental condition and task length, \( F(3, 21) = 13.83, p < 0.0001, \eta^2 = 0.66, \) and by the results of the contrast tests. For the SC condition, Bonferroni \( t \) tests showed significant differences between Level 1 and Level 4 and between Level 2 and Level 4, but no significant difference was found between Levels 3 and 4. For the SA condition, Bonferroni \( t \) tests revealed no significant differences in speech naturalness among the four levels of fingerspelling task length.

**Relationship of Acoustic Measures to Speech Naturalness**

In order to examine the relationship between the speech production and perception parameters analyzed in the ANOVAs reported above, a multiple regression analysis was performed using the five temporal characteristics of speech as predictor variables and the scaled speech naturalness as the predicted variable. Table 3 displays the bivariate correlations among the five temporal characteristics and speech naturalness. As expected, sentence duration showed high correlations with the other four temporal measures (ranging from 0.73 to 0.84) because total sentence duration is the sum of all durational elements in the sentence. The correlations among the other temporal measures were low to moderate (ranging from 0.25 to 0.79), indicating that these variables were good candidates for entry into a regression analysis as predictor variables that could complement each other in the prediction of speech naturalness. The multiple \( R \) using all five predictors that entered significantly into the regression equation was \( R = 0.92, \) with an \( R^2 \) of 0.84 and a standard error of estimate of 1.08 on the 9-point naturalness scale. As indicated by the correlations in Table 3, the order of the independent variables entered into the regression equation for predicting speech naturalness was (1) sentence duration, (2) diphthong duration, (3) interword interval before the experimental word, (4) word duration, and (5) interword interval after the experimental word. The results of this regression analysis demonstrate the significant predictive relationship of these selected temporal characteristics of speech production and speech naturalness.

**Discussion**

These results demonstrate a consistent pattern of significantly increased sentence, word, diphthong, and interword interval durations as speakers move from the SA to the SC condition. In addition, sentence, word, and interword interval before the experimental word all showed significant increases with increased fingerspelling task length. The present experimental findings are consistent with previous results that have shown elongated temporal speech characteristics when signed English and fingerspelling are combined with speech in SC by both inexperienced signers (Whitehead, Schiavetti, Metz, & Farrell, 1999) and experienced signers (Whitehead et al., 1995). Attempting to speak and use manual communication at the same time impacts a speaker’s speech timing, regardless of the type of manual system used in SC (Schiavetti, Whitehead, & Metz, 2004; Windsor & Fristoe, 1989, 1991). Thus, because SC attempts to coordinate the rapid speech production act with the slower sign production act, combining the two communication modes results in a slowing of the more rapid speech act to maintain simultaneity with the slower signing act (Windsor & Fristoe, 1989, 1991).

Although the pattern of increased durations by the inexperienced signers was very similar to that shown by experienced signers, the absolute values of the duration increases were much larger for the inexperienced signers, who showed greater mean lengthening of sentences, words, and interword intervals before and after the experimental word to be fingerspelled than the experienced signers did during SC on the same task (Schiavetti et al., 1998). The inexperienced signers also showed larger standard deviations in these measures than the experienced signers, except for diphthong duration. These findings, then, indicate not only greater overall temporal elongation of segments and pauses for inexperienced signers, but also more variability among inexperienced than experienced signers.

To further substantiate the observations of larger SA–SC differences among the inexperienced signers, a meta-analysis was undertaken to compare the effects of SA versus SC on each of the dependent variables for the inexperienced and experienced signers using Cohen’s \( d \) (Schiavetti & Metz, 2006). An effect size was calculated separately for each group of participants for the SA–SC communication mode difference on each dependent variable using Cohen’s \( d \) for these bivalent comparisons with the conservative approach of pooling the standard deviation across groups (Cohen, 1988). Table 4 displays the effect sizes for the inexperienced and experienced signers for each dependent variable.

Inspection of Table 4 reveals that the inexperienced signers’ SA–SC effect size for sentence duration was almost double the experienced signers’ SA–SC effect size, indicating that the inexperienced signers slowed their speech much more than the experienced signers did. The effect sizes for word duration and interword interval after the fingerspelled

<table>
<thead>
<tr>
<th>Variable</th>
<th>SD</th>
<th>WD</th>
<th>DD</th>
<th>IWIB</th>
<th>IWIA</th>
<th>SN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence duration (SD)</td>
<td></td>
<td></td>
<td>.81</td>
<td></td>
<td></td>
<td>.86</td>
</tr>
<tr>
<td>Word duration (WD)</td>
<td></td>
<td></td>
<td>.77</td>
<td></td>
<td>.84</td>
<td>.73</td>
</tr>
<tr>
<td>Diphthong duration (DD)</td>
<td></td>
<td></td>
<td>.61</td>
<td></td>
<td>.79</td>
<td>.59</td>
</tr>
<tr>
<td>Interword interval before word (IWIB)</td>
<td></td>
<td></td>
<td>.51</td>
<td></td>
<td>.79</td>
<td>.59</td>
</tr>
<tr>
<td>Interword interval after word (IWIA)</td>
<td></td>
<td></td>
<td>.79</td>
<td></td>
<td>.79</td>
<td>.59</td>
</tr>
<tr>
<td>Speech naturalness (SN)</td>
<td></td>
<td></td>
<td>.60</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
word were similar for the inexperienced and experienced signers and are consistent with the conclusion of Windsor and Fristoe (1989, 1991) that time is needed after an experimental word to equalize the time difference between the manual and speech tasks needed to produce the word.

Inspection of Table 4 also reveals that the effect sizes for diphthong duration and interword interval before the target word were much larger for the inexperienced signers than for the experienced signers, accounting for most of the increase in sentence duration. These effect sizes for the dependent variables demonstrate that the inexperienced signers showed greater duration increases just before the fingerspelled word than the experienced signers. This indicates a possible need for more planning time for fingerspelling for these inexperienced signers than for experienced signers in anticipation of the increasingly difficult fingerspelling task. Indeed, this notion is consistent with our anecdotal observation that some of the inexperienced signers fingerspelled the target word before speaking it and suggests that future research employing video recording of inexperienced signers would be important in determining their strategies for coping with the conflicting temporal demands of speech and manual communication during SC.

The performance difference between inexperienced and experienced signers in anticipation of the experimental word to be fingerspelled was also interesting with respect to the difference between the elongation patterns of the diphthong and the interval before the fingerspelled word. In both variables, inexperienced signers showed more than double the elongation of the experienced signers, but inexperienced signers showed much less variability in diphthong duration and much more variability in interword interval variability than the experienced signers. These results indicate that all of the inexperienced signers were slowing down their speech more on the diphthong in anticipation of the fingerspelling task, but that some of them were taking much longer pauses before the experimental fingerspelled word. Possibly those were the participants who were fingerspelling during the pause rather than during the word itself. Perhaps this was a strategy to catch fingerspelling up with the spoken word as a method of salvaging simultaneity. The inexperienced signers’ greater uniformity in lengthening the diphthong duration and their more varied approach to the interword interval in anticipation of the fingerspelling tasks again agrees with our anecdotal evidence that some inexperienced signers paused before the target word and fingerspelled the word in silence and then said the word. These anecdotal observations need to be explored in future research with video recording of the signers in order to confirm the possibility that at least some inexperienced signers are fingerspelling during pause time rather than during the actual fingerspelled word.

If this pattern of fingerspelling during pause time is confirmed, it may shed light on the learning process of both fingerspelling and SC. It is interesting to note that Akamatsu and Stewart (1989, p. 370) found that “hearing people learning to sign report that it is very difficult to learn fingerspelling” and Wilcox (1992) found that learning to read fingerspelling was rated as the most difficult learning task by hearing adults who were enrolled in a sign language course. Akamatsu and Stewart also noted that the demanding task of coordinating manual orthographic patterns and articulatory speech patterns may present sequential conflicts for SC users that further increase the difficulty of fingerspelling and speaking simultaneously. Thus it would not be surprising if inexperienced signers need to fingerspell during pause time as they learn to use SC, but would later fingerspell more rapidly as they become more fluent in both manual communication and SC. It is possible that as inexperienced signers mature into experienced signers, their increased rapidity of fingerspelling develops into increased simultaneity with the spoken word in SC. Critical issues for future research include learning (a) the amount of time and practice that are necessary to become fluent enough in SC to fingerspell and speak simultaneously and (b) what strategies might be used by beginning signers to develop fluency in SC. Video recording of both inexperienced and experienced signers’ performance on these tasks would be important in further investigations of the effects of fingerspelling task difficulty on speech characteristics during SC to provide better generalization to the demands of real-world simultaneous oral and manual communication.

The results regarding speech naturalness and its relationship to temporal characteristics of speech for the inexperienced signers were quite similar to those for the experienced signers (Schiavetti et al., 1998). SC speech was perceived as more unnatural than SA, and naturalness decreased systematically with fingerspelling task length in the SC mode, but not in the SA mode. Scaled speech naturalness could be predicted reasonably well from the independent variables that demonstrated various temporal elongations that accompany SC. Increases in sentence, word, diphthong, and interword interval duration all correlated with speech naturalness and entered into the regression equation for its prediction in a similar manner to the equation reported by Schiavetti et al. (1998) for the experienced signers with one exception. The diphthong duration and interword interval before the fingerspelled word were the second and third predictor variables behind sentence duration for our inexperienced signers, whereas experimental word duration and interword interval after the fingerspelled word were the second and third predictors for the experienced signers. These results are consistent with the notion that inexperienced signers differ from experienced signers in their anticipation of the fingerspelling task and may be using excessive pause time to fingerspell rather

Table 4. Effect sizes (Cohen’s d) for comparisons of speech alone versus simultaneous communication for inexperienced and experienced signers on sentence, word, diphthong, interword interval before, and interword interval after durations, and speech naturalness.

<table>
<thead>
<tr>
<th></th>
<th>Inexperienced signers</th>
<th>Experienced signers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence duration</td>
<td>6.30</td>
<td>3.54</td>
</tr>
<tr>
<td>Word duration</td>
<td>2.95</td>
<td>2.72</td>
</tr>
<tr>
<td>Diphthong duration</td>
<td>3.95</td>
<td>1.39</td>
</tr>
<tr>
<td>Interword interval before</td>
<td>4.89</td>
<td>1.85</td>
</tr>
<tr>
<td>Interword interval after</td>
<td>1.81</td>
<td>1.74</td>
</tr>
<tr>
<td>Speech naturalness</td>
<td>4.94</td>
<td>2.69</td>
</tr>
</tbody>
</table>
than actually simultaneously fingerspelling spoken words. Again, future research with video recording may help to confirm this speculation regarding the development of SC in beginning signers.

Hyde, Power, and Leigh (1998) discussed the possibility of a trading relationship between perceived speech naturalness and intelligibility of SC. Because their work was with experienced signers, this possibility needs to be addressed with inexperienced signers as well in order to gain a better understanding of the efficacy of SC as a speech model. It is especially important to compare the intelligibility of inexperienced and experienced signers because the experienced signers demonstrated greater speech naturalness in SC than did the inexperienced signers.

Future research concerning SC with inexperienced signers needs to take a longitudinal approach to study the development of sign and fingerspelling skills as they are integrated with speech in SC. Of great importance is the issue of how family members and other significant persons in the language learning environments of deaf and hard-of-hearing children are learning manual communication and what training strategies are most effective for helping them to communicate more efficiently through SC as their manual skills develop. Questions concerning the rate of manual skill development, learning strategies employed by new signers, and teaching strategies used by sign language teachers and the influence of these factors on the quality of speech, sign, and fingerspelling are very complex issues that will require multivariate research designs. Such data are important for the development of recommendations for the ongoing learning and use of SC by persons with newly acquired manual communication skills. These recommendations should enhance the communicative effectiveness for all members within the deaf or hard-of-hearing child’s environment.

ACKNOWLEDGMENT

A portion of this research was undertaken at the National Technical Institute for the Deaf in the course of an agreement between the Rochester Institute of Technology and the U.S. Department of Education. Part of this research was supported by funds from the Geneseo Foundation provided through the Research Council.

REFERENCES


Contact author: Robert Whitehead, National Technical Institute for the Deaf, Rochester Institute of Technology, Carey 1309, 96 Lomb Memorial Drive, Rochester, NY 14623-5604. E-mail: rw-wncr@rit.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License.