Acoustically Derived Perceptual Evidence for Coarticulatory Errors in Apraxic and Conduction Aphasic Speech Production

Malcolm R. McNeil, Michiko Hashi, and Helen Southwood

INTRODUCTION

Deficits in the temporal and spatial coordination of speech have been suggested by many (Itoh, Sasanuma, & Ushijima, 1979; Kent & Rosenbek, 1983; Rosenbek, Kent, & LaPointe, 1984; Wertz, LaPointe, & Rosenbek, 1984) to explain the mechanisms underpinning apraxia of speech. Data have been amassed from acoustic, physiologic, and perceptual studies to support these suggestions and to attribute the observed speech deficits to a phonetic/motoric level of the speech production system (see McNeil & Kent, 1990, for a review of this evidence). Recent evidence from acoustic (Kent & McNeil, 1987; McNeil, Liss, Tseng, & Kent, 1990), kinematic (McNeil & Adams, 1991; McNeil, Weismer, Adams, & Mulligan, 1990; McNeil, Hashi, & Tseng, 1991) and perceptual (Odell, McNeil, Rosenbek, & Hunter, 1990, 1991) studies have also supported the idea that at least some individuals with conduction aphasia may also present phonetic/motoric deficits in addition to, or instead of, their assumed phonologic-level speech errors.

Among the more salient sources of evidence for temporal and spatial coordination speech production deficits in the apraxia of speech population are data provided from a perceptual approach used first with this population by Ziegler and Von Cramon (1985). In this paradigm, normal listeners judge the presence of a particular vowel in CV or CCV sequences when the vowel has been systematically reduced (computer edited) in a right-to-left fashion. This approach eliminates portions of the sound rang-
ing from only a few milliseconds to the entire vowel. In the Ziegler and Von Cramon study, utterances from normal and apraxic subjects’ productions were judged by normal listeners. These judges were required to identify or predict the vowel that followed each stimulus. Each utterance was “gated” at five different durations, and the listener selected responses from three vowels. The premise underlying this procedure is that, even if a listener hears only part of a vowel—or in some cases, none of it—the consonant carries sufficient information about lip-retracted (/i/) or -protruded (/U/) vowels to predict the vowel that follows. This premise is supported by acoustic analyses of normal speech (Ohman, 1966; Sereno, Baum, Marean, & Lieberman, 1986; Soli, 1981) wherein a spectral peak in the region of the second formant of the following vowel has been identified in the consonant portion of the CV sequence that is believed to carry the cue for the vowel. This assumes that in normal speakers, the speech plan (the selection and ordering of the phonological units) and the motor speech program (the conversion of the phonological information into instructions for movement) prepare the articulators in advance of their execution. Further assumed is that an impaired speech motor program would be indicated by the production of consonants that do not carry sufficient information about the correctly selected vowel for their accurate prediction.\(^1\) In such a task, the amount of acoustic information available to the listener is a function of the gating, and the accuracy of vowel identification represents the efficiency of coarticulation as well as the integrity of the motor speech programmer (assuming that the speech plan has been assembled properly). Results for Ziegler and Von Cramon’s normal speakers confirmed that only the lip retracted or protruded consonantal acoustic information, and little or none of the vowel, was needed to accurately predict the vowel. Results for two apraxic speakers, however, revealed that more of the vowel was needed for correct identification from their productions. This result was interpreted as a reflection of a delayed onset of anticipatory vowel gestures and an impaired motor speech programmer in the apraxic speakers.


\(^1\) Sussman, Marquardt, MacNeilage, and Hutchinson (1988) point out that analyses restricted to “on-target” productions omit important data and sources of evidence that are available with the analysis of error data. However, in the present study, it is imperative that the experimenters’ target vowel was also the subject’s target vowel; the source of the deficit in the speech production process thus must be attributed to the motor-programming level rather than to the speech-planning level (a level often implicated in many aphasic phonologic errors and in normal “slips of the tongue.”)
"posterior," "fluent," and "nonfluently"—and normal subjects by Katz (1987), Katz, Machetanz, Orth, and Schonle (1990a), and Tuller and Story (1987) revealed inconsistent patterns of anticipatory lingual and labial coarticulation for the aphasic populations compared to the normal control subjects. A kinematic study of anticipatory coarticulation in two anterior aphasic subjects (Katz, Machetanz, Orth, and Schonle, 1990b) revealed coarticulatory patterns that were more variable than the same patterns of two control subjects. This variability was found in displacement and not in the temporal aspects of the movement.

Because of the inconsistent results across both perceptual and acoustic studies of anticipatory coarticulation, the present study was designed as a replication in apraxic speakers and an extension of these analyses to the speech of conduction aphasic subjects. This extension is of particular interest because speech problems in conduction aphasia have traditionally been assigned exclusively to the phonological level of speech production. Although some challenges to this assignment have been proposed (Kent & McNeil, 1987; McNeil, Liss, Tseng, & Kent, 1990; McNeil, Hashi, & Tseng, 1991; Tseng, McNeil, Adams, & Weismer, 1990), there is insufficient evidence to argue conclusively for a phonetic assembly, motor-programming, or execution mechanism deficit in this population.

METHOD

Subjects

The speech samples for this study were taken from one normal speaker, two subjects with apraxia of speech uncomplicated by dysarthria or aphasia, and two subjects diagnosed as having conduction aphasia without concomitant apraxia of speech or dysarthria. The selection criteria for all subjects have been summarized elsewhere (e.g., McNeil & Adams, 1991) and are not repeated here. These subjects were chosen from the pool of subjects used in a series of studies, and their biographic characteristics, speech, language, and cognitive performance on a variety of measures are summarized in Table 1.

Six normal adults served as judges for this study. They were considered sophisticated listeners in that they: (1) had successfully completed a college-level course in phonetics and a graduate-level course in motor speech disorder; (2) had at least 1 year of research or clinical experience in neuro-

---

2. The subjects used in this investigation have been used in a series of studies. The following scheme identifies the subjects used in both the present study (first designation) and the original one (second designation): \( N_1 = N_1, A_1 = A_3, A_2 = A_2, C_1 = C_4, C_2 = C_3 \).
Table 1. Summary of Biographical and Descriptive Data for Normal Control, Apraxic, and Conduction Aphasic Speakers

<table>
<thead>
<tr>
<th>Subject Characteristics</th>
<th>N₁</th>
<th>A₁</th>
<th>A₂</th>
<th>C₁</th>
<th>C₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>67</td>
<td>54</td>
<td>62</td>
<td>62</td>
<td>60</td>
</tr>
<tr>
<td>Structural-functional exam. ¹</td>
<td>WNL</td>
<td>WNL</td>
<td>WNL</td>
<td>WNL</td>
<td>WNL</td>
</tr>
<tr>
<td>Total RCPM ²</td>
<td>32</td>
<td>30</td>
<td>28</td>
<td>27</td>
<td>32</td>
</tr>
<tr>
<td>Overall RTT ⁴</td>
<td>14.35</td>
<td>12.23</td>
<td>12.08</td>
<td>13.94</td>
<td>13.04</td>
</tr>
<tr>
<td>BDAE Aud. Comp. ⁵</td>
<td>119</td>
<td>118</td>
<td>113</td>
<td>114</td>
<td>117</td>
</tr>
<tr>
<td>BDAE speech rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artistic agility</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Phrase length</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Melodic line</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>BDAE total sent. repetitions w/o errors</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Apraxia battery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total limb</td>
<td>48</td>
<td>48</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Total oral</td>
<td>50</td>
<td>49</td>
<td>47</td>
<td>49</td>
<td>49</td>
</tr>
</tbody>
</table>

⁴S-F Exam. Structural-Functional Examination.
⁵WNL There was a questionable right-sided lingual weakness on clinical examination that was not confirmed with additional testing for this subject.
Total RCPM Total number correct on the Raven Coloured Progressive Matrices.
Total WFM Total Word Fluency Measure score.
O.A. PICA Overall score on the Porch Index of Communicative Ability (see text for method of calculation).
O.A. RTT Overall score on the Revised Token Test.
BDAE Aud. Comp. total number correct on all four auditory comprehension subtests of the Boston Diagnostic Aphasia Examination.
BDAE Spch. Rtg. Ratings assigned for articulatory agility, phrase length, and melodic line from the rating of speech characteristics section of the Boston Diagnostic Aphasia Examination.
BDAE Total Sent. Rep. w/o Errors Total number of sentences repeated without articulatory errors from the sentence repetition subtest of the Boston Diagnostic Aphasia Examination.
Apraxia Bat. for Adults Total score on the limb and oral apraxia subtests from the Apraxia Battery for Adults. (A score of 50 represents error-free performance.)
WNL There was a questionable oral sensory diminution on clinical examination in this subject.
genic communication disorders or in developmental phonological disorders; (3) had no speech or language problems; (4) were native speakers of American English; and (5) had normal hearing by self-report.

**Stimuli**

The single normal and four pathologic speakers' productions of 10 single words (*six, sixteen, stop, stopping, build, building, big, bigger, bob, and bobby*), with either an /S/, /ST/, or /B/ consonant preceding either the lip-retracted vowel /I/ (as in *six*) or the lip neutral vowel /a/ (as in *Bob*), were used as stimuli. Only utterances judged as *on target* at the word level, using a broad phonetic transcription reference, were employed as stimuli. This selection of data allowed the elimination of obvious phonemic-level errors and increased the interpretability of the data relative to the assignment of phonetic/motoric mechanisms for any observed effects. Five repetitions of each utterance from each speaker were digitized at a 20-kHz sampling rate using a 10-kHz low-pass filter. The stimuli from each speaker were computer edited (gated) at five different durations and were randomized and rerecorded onto audiotape. Figure 1 shows the acoustic waveform for one utterance (*six*) used in this study. The five different gates that were presented to each normal judge are shown. Gate 5 represents the consonant plus the entire vowel and should be identified correctly, beyond chance-level errors, in all instances of correct production.

**Judge's Task**

The judges listened to all 250 stimuli and selected, from among five choices (/a/, /i/, /ae/, and /ə—the schwa), the vowel that followed each consonant. A probability of identifying the correct vowel by chance was 20%, or 1 in 5, if the judges believed that all five vowel choices were present (though only the /I/ and /a/ vowels were present in the stimuli). All listening sessions were conducted in a sound-treated chamber, the tapes were presented in a sound field, and the listeners selected the volume with which they felt most comfortable. Listeners were required, and frequently encouraged, to guess when they were not certain.

**Analysis**

All responses were tabulated by speaker and utterance. Binomial probabilities were calculated for each combination of subject, utterance, and
Figure 1. Acoustic waveform for the utterance “six” showing complete signal for gate 5, with successive portions of the signal removed (gaited) to gate 1, where only the /s/ was retained for perceptual judgment.

gate. The alpha level was set at .05 for each comparison and was calculated as a one-tailed test.

RESULTS AND DISCUSSION

The general results of this investigation are illustrated in Figures 2 through 11. These figures represent the data for each subject for each of the 10 stimuli used in this study and show the probability at which the correct vowel was identified. The horizontal line at the bottom of each figure represents the .05 confidence level. Symbols falling below this level represent instances in which the vowel was identified correctly at a rate significantly beyond chance level for that subject on that particular utterance. Symbols falling above this line represent vowels that were judged incorrectly or exceeded the .05 confidence level for that particular subject. The ordinate in these figures, then, represents the alpha level achieved for that subject on that particular stimulus at each gate. Gate 5 contained the consonant plus the entire vowel and should in all instances be identified correctly beyond chance level for all subjects whose productions were perceived accurately by these judges. There were instances in which the stimulus at Gate 5 was not identified significantly beyond chance level.
Figure 2. Confidence level achieved for each individual subject's judgments at each of the five gaited stimuli for the normal and pathologic subject's productions of the word "STOP." Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.

Figure 3. Confidence level achieved for each individual subject's judgments at each of the five gaited stimuli for the normal and pathologic subject's productions of the word "STOPPING." Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.
(as was the case with subjects C₂ and A₁ for the utterance six; see Figure 4). In these cases, it is difficult to interpret the data as evidence that the vowels identified at earlier gates represent coarticulatory effects rather than the selection of the incorrect vowel and, hence, a speech-planning (i.e., phonologic-level) error for these particular subjects on that particular utterance. However, in some instances, as is the case with subject A₁ in the same trial, the vowel was identified correctly beyond chance levels at all gates preceding Gate 5. It might be argued that alterations in the vowel itself caused the listener to perceive abnormalities when the entire vowel was produced; consequently, an incorrect vowel was identified with an increase in acoustic information.

Figure 6 (the utterance build) shows that the vowel productions for the N₁ and C₁ were judged accurately at all gates. However, A₁ and C₂ were judged inaccurately for Gates 1 through 3 and 1 through 4, respectively. It should be noted, however, that because A₂ was judged inaccurately at Gate 5, the result is difficult to interpret as a coarticulatory error for this particular subject and utterance.

Figure 7 shows the results for the utterance building. As with the utterance build, the vowels for N₁ and C₁ were judged accurately at all gates (save Gate 2 for N₁), whereas A₂ and C₂'s vowels were not identified accurately until considerably more of the vowel was present. With the correct identification of the vowel when it was acoustically present, the differences can more readily be assigned to the effects of coarticulatory differences in the speech of the two apraxic and one conduction aphasic subjects. Results for the remainder of the utterances are shown in Figures 8 through 11. In general, they paralleled those for the previous two utterances.

Overall, it was found that the neutral vowel /a/ in stop and stopping did not provide sufficient acoustic cues in the consonant for any of the subjects (including the normal subject) to predict the vowel above chance level until all or nearly all of the vowel was heard. This is interpreted as support for the notion that there are coarticulatory cues imbedded in the consonantal portion of the signal for some phonetic contexts, such as lip-retracted vowels, and that this experiment was in line with the experimental paradigm.

With the exception of those in the utterances stop and stopping, C₁'s vowels were identified correctly at a rate significantly beyond chance for all words at all gates. Likewise, the normal subject's vowels were generally identified correctly at all gates for all utterances, and in all instances in Gates 4 and 5. A₁, A₂, and C₂ all demonstrated productions in which the vowel could not be identified correctly beyond chance level for any of the utterances (though not at every gate for all utterances).

These results are of interest in two ways. First, the coarticulation deficits in apraxia of speech demonstrated by Ziegler and Von Cramon (1985) were replicated in the present study, though with less consistency than
Figure 4. Confidence level achieved for each individual subject's judgments at each of the five gaited stimuli for the normal and pathologic subject's productions of the word "SIX." Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.

Figure 5. Confidence level achieved for each individual subject's judgments at each of the five gaited stimuli for the normal and pathologic subject's productions of the word "SIXTEEN." Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.
Figure 6. Confidence level achieved for each individual subject’s judgments at each of the five gaited stimuli for the normal and pathologic subject’s productions of the word “BUILD.” Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.

Figure 7. Confidence level achieved for each individual subject’s judgments at each of the five gaited stimuli for the normal and pathologic subject’s productions of the word “BUILDING.” Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.
Figure 8. Confidence level achieved for each individual subject’s judgments at each of the five gaited stimuli for the normal and pathologic subject’s productions of the word “BIG.” Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.

Figure 9. Confidence level achieved for each individual subject’s judgments at each of the five gaited stimuli for the normal and pathologic subject’s productions of the word “BIGGER.” Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.
**Figure 10.** Confidence level achieved for each individual subject's judgments at each of the five gaited stimuli for the normal and pathologic subject's productions of the word "BOB." Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.

---

**Figure 11.** Confidence level achieved for each individual subject's judgments at each of the five gaited stimuli for the normal and pathologic subject's productions of the word "BOBBY." Judgments falling below the horizontal line in the figure at the .05 alpha level (probability of correct vowel identification) represent significantly correct vowel identification. Symbols falling above this line represent nonsignificant vowel identification.
reported by those authors. Second, the results demonstrated abnormalities of coarticulation in one of the two conduction aphasic speakers. Vowels from C₁ were identified correctly to the same extent as those from the normal speaker. This is interpreted to imply that this admittedly limited test of anticipatory coarticulation did not show C₁ to have problems in that area. Conversely, vowels from A₂ were identified no better than those from the apraxic speakers, suggesting the existence of coarticulatory problems in this particular subject. Taken together, the results suggest that conduction aphasic speakers may well vary in their speech production ability, and some of them may demonstrate anticipatory coarticulatory deficits similar to apraxia of speech.

It might be speculated that the anticipatory coarticulatory deficits found in this conduction aphasic speaker are in some way related to lesion location. C₁, who demonstrated normal coarticulation, had a widespread lesion that included part of Broca's area; pre- and postcentral gyri; the superior parietal, supramarginal, and angular gyri; Wernicke's area; part of middle temporal gyrus; and part of the insula. In contrast, C₂, who demonstrated coarticulatory deficits similar to the apraxic subjects, had a lesion in the supramarginal, part of superior parietal, and angular gyri; Wernicke's area; part of middle temporal gyrus; and part of the insula. No lesion was found in either the precentral gyrus or in Broca's area in this conduction aphasic subject. Such clinicoanatomical relations contradict the subclassification of conduction aphasia proposed by several aphasiologists such as Kertesz (1979), suggesting instead a motor component in conduction aphasic subjects with anterior lesions. They also call to question the classification systems used in previous anticipatory coarticulation studies, such as the anterior/posterior dichotomy and perhaps the fluent/nonfluent dichotomy, for neither of these two conduction aphasic subjects would have been classified as nonfluent aphasic subjects using the criteria of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983). This result is also difficult to reconcile with disconnection and phonologic accounts for the speech of conduction aphasic subjects and suggests a more complex explanation for the speech production errors of these subjects than such models provide. The findings also highlight the inadequacy of the classification system employing conduction aphasia for representing the underlying nature (linguistic vs. motoric) for some of the deficits on which the classification is based (i.e., phonologic speech errors that are more frequent on repetition than in spontaneous speech or in reading).

Adding to this classification dilemma, Valdois, Joanette, Nespoulous, and Poncet (1988) have questioned whether some subjects typically classified as conduction aphasic are not actually afferent motor aphasic subjects or, alternatively, whether afferent motor aphasia is not a subform of conduction aphasia. Given the lack of established validity for this cate-
gory in any classification system, the lack of established criteria for selecting afferent motor aphasic subjects, and the lack of knowledge as to whether linguistic or motoric mechanisms underlie the speech production errors in this population, these alternatives add little to a clinically useful application or theoretical explanation for the behaviors observed in the conduction aphasic subjects in this investigation.

The lack of anticipatory coarticulation effects in the conduction aphasic subject whose performance matched the normal subject’s could be explained as a matter of severity of deficit rather than as a difference in the nature of the mechanisms generating the errors or the lack thereof. The problem with this potential explanation is that there is no clear, unbiased (with respect to level of analysis), and preferred method for determining the severity of the speech production deficit. In spite of these problems, the subjects from this study are in agreement with this hypothesis. The conduction aphasic subject who demonstrated no evidence of an anticipatory coarticulatory deficit, C1, repeated fewer sentences without errors and had a higher score on the melodic line rating of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983) and less aphasia as measured by the Porch Index of Communicative Ability (Porch, 1967) and the Revised Token Test (McNeil & Prescott, 1978) than C2, who did demonstrate an anticipatory coarticulation deficit. This extremely small sample size does not allow a conclusion from this observation; however, there may be reason to formulate and test this hypothesis. A confirmation of this notion has implications that reach far beyond the mere involvement of a motor speech mechanism in some or all of the speech errors of apraxic and conduction aphasic individuals.

The data from this investigation are consistent with the attribution to at least some conduction aphasic subjects a motor-speech-level deficit responsible for at least some of their errors. These findings also are consistent with others demonstrated by perceptual, acoustic, and physiologic studies by our research group, although we obviously need additional subjects in this and other analyses before we can generalize these findings. It is clear, however, that the diagnosis of conduction aphasia does not require the speech production errors to be a matter of forming slots, selecting phonologic elements, or filling the slots with phonologic information that has been properly selected and sequenced (in the model of Shattuck-Hufnagel, 1979, 1983, 1987).

REFERENCES


