RELATIONSHIPS BETWEEN ORAL SENSORY FEEDBACK SKILLS AND ADAPTATION TO DELAYED AUDITORY FEEDBACK

JOSEPH S. ATTANASIO
Department of Communication Sciences and Disorders, Montclair State College

The relationship between adaptation to delayed auditory feedback (DAF) and oral sensory feedback as measured by tasks of oral form discrimination (OFD) was investigated. In addition, the interactions between OFD ability and attempts to heighten subjects' awareness of oral feedback while speaking under DAF conditions by giving instruction to subjects as well as the condition of instruction itself were studied with adaptation to DAF as the dependent variable. Analyses of the data indicated a relationship between OFD ability and adaptation to DAF and an interaction between OFD ability and instruction. The study's rationale and results are discussed within the context of servosystem theory.

INTRODUCTION

A prominent consideration of research workers in speech science and speech-language pathology has been the attempt to describe and explain the manner in which the human speech system operates. "System" is used here to refer to the interacting and interdependent components of a functional unit that is only partially accessible to direct observation. Because the speech system's complexity does not allow completely direct observation of its operations, models have been developed by a number of investigators as a means of representing what might be the nature of the system and how it functions (Lee, 1950; Lashley, 1951; Fairbanks, 1954; Liberman et al., 1962, 1967; Eilenberg, 1973; Kent, 1973).

Servosystem Models

Many models have employed the concepts of servosystem theory. Though servosystem models differ from one another along several dimensions, they all share the view that speech production is monitored and controlled by feedback of afferent stimuli via sensory information received from...
peripheral receptors (Kramer, 1972; Ringel, 1970). Sensory feedback, as described in servosystem models, is provided by the auditory, oral tactile, and oral kinesthetic–proprioceptive pathways (Fairbanks, 1954; Kramer, 1972; MacNeilage, 1970).

A number of studies have attempted to investigate the importance of feedback control by studying what happens to speech when sensory feedback is altered or manipulated. For example, the literature describing the effects of delayed auditory feedback (DAF) on speech indicates that, when individuals speak under DAF, speech is disrupted along one or more of the dimensions of time, rate, fluency, and articulation (Fairbanks, 1955; Fairbanks and Guttman, 1958; Lee, 1950; Webster and Dorman, 1971). Often, the finding of speech disruption is interpreted as confirmation of the need for auditory feedback to monitor and control speech production and as validation of the servosystem theory in general.

While studies in auditory feedback have been used to support the validity of servosystem theory, most models ascribe some function to oral tactile and oral kinesthetic–proprioceptive feedback (Ringel, 1970; Kramer, 1972). The research that is available in oral tactile and oral kinesthetic–proprioceptive feedback, whether it be speculative or empirical, suggests that these two feedback mechanisms may be at least as important as auditory feedback in guiding speech production (Van Riper and Irwin, 1958:110; Henke, 1970; MacNeilage, 1970; Perkell, 1969).

Oral Sensation

Research concerned with oral sensation has indicated that oral tactile and oral kinesthetic–proprioceptive feedback deficiencies can disturb speech production, especially articulation. Studies by Wilson (1960) and Ringel (1970), for example, have included observations of individuals with sensory pathologies and descriptions of speech production in persons who have experienced experimentally induced oral sensory deficiencies. It would appear that servosystem control of speech is implicated since oral feedback does not function adequately in such cases, thereby contributing to disturbances in speech production.

The view, then, on the part of several researchers mentioned above who have studied servosystem theory is that speech production is monitored and controlled by auditory, oral tactile, and oral kinesthetic–proprioceptive feedback and that disturbances in any of these sensory channels can impair speech.

Auditory and Oral Sensory Interactions

With the development of instruments that can produce DAF, researchers have been able to stimulate stutteringlike speech behaviors and adaptation of those behaviors in normal speakers within the context of investigating
the role that auditory feedback plays in speech monitoring and control (e.g., Atkinson, 1953; Black, 1951; Burke, 1975; Fairbanks, 1955; Kramer, 1972; Lee, 1951; Tiffany and Hanley, 1956). Adaptation to DAF is the decrease, over successive oral readings of the same passage under DAF conditions, of the nonfluencies, articulation errors, and other speech errors that result from the disruptive effects of DAF on speech. Essentially, adaptation to DAF refers to improvement in initial levels of speech performance under DAF over time and with practice (Attanasio, 1977; Venkatagiri, 1980). Some DAF research has attempted to explain or describe the adaptation phenomenon as a reflection of reduced reliance on auditory feedback with a concomitant shift to other types of sensory feedback for speech monitoring. These shifts are made through a variety of conscious or unconscious strategies employed by speakers under DAF (Gruber, 1965; Rouse and Tucker, 1966; Webster and Dorman, 1971; Webster and Lubker, 1968; Wilson, 1960). Considered together, these studies can be interpreted as suggesting that during DAF conditions speakers shift to nonauditory feedback (oral tactile and oral kinesthetic–proprioceptive feedback from the articulators) in order to override the effects that aberrant auditory feedback has on speech.

A study by Kramer (1972) highlights the possible interactions between auditory and oral feedback. Kramer investigated the relationships between speech disruption under one instance of DAF and oral form discrimination (a task taken to be a measure of oral sensory feedback). Forty subjects ranked high in oral form discrimination ability and 40 subjects ranked low in oral form discrimination ability read a passage aloud under a 0.20-second delay in auditory feedback. Results showed that there was a statistically significant probability of a difference between the two groups in the total number of pooled temporal errors made under DAF. The low oral form discrimination ability group made more errors of word or syllable repetition plus sound or syllable prolongation than did the group with high oral form discrimination ability.

Purpose of the Study

The present study was undertaken to investigate the relationships between oral and auditory feedback as they function in speech production. The study may be viewed as an attempt to extend the research in the application of servosystem theory to speech. The design of the study and its rationale may be summarized as follows: (1) Delayed auditory feedback disrupts normal auditory monitoring of speech production, thereby causing articulatory errors and nonfluencies; (2) nondisturbed, nonauditory feedback (oral tactile and oral kinesthetic–proprioceptive feedback from the articulators) may be used by speakers over time to reduce the effects of delayed auditory feedback; and (3) the ability to use oral tactile and
oral kinesthetic–proprioceptive feedback depends upon a speaker's sensitivity to that feedback and upon his ability to shift from auditory to oral monitoring of speech production. Two methods were combined to investigate speakers' sensitivity to oral feedback. One method involved oral form discrimination tasks as measures of an intrinsic awareness of oral sensation or oral feedback, and the other method, based on the literature in artificial knowledge of results (Holding, 1970), attempted to heighten sensitivity through extrinsic means by instructing subjects to attend to fluency and to articulatory movement and articulatory position while reading aloud under DAF conditions. Artificial knowledge of results (KR) is defined as feedback of information to a subject by a source external to the subject concerning the results of his actions (Holding, 1970). Holding stated that artificial KR may be useful in alerting subjects to such intrinsic cues as kinesthetic sensitivity, of which there is usually little awareness.

METHOD

Subjects

The sample consisted of 40 subjects, 34 of whom were female and 6 male. They ranged from 18 to 25 years of age, with a mean of 20.6 years. The subjects volunteered for the study by responding to advertisements, announcements, and notices posted throughout the campus of the college that they were attending. Subjects had no prior experience relative to the purposes of the study.

The sample was limited to individuals who were normal speakers with normal oral speech mechanisms and normal hearing. They were free from past or present neuromotor or neurosensory problems, verified through case histories and interviews with the subjects. Health status was confirmed by an examination of each subject's school medical records after the permission of the subject was obtained.

Equipment and Procedures for Phase I of the Study—DAF

Auditory sensory feedback was experimentally disrupted by the use of DAF. A Viking 87 Tape Transport recorder associated with an Allison model 22 Clinical and Research Audiometer was used to deliver the DAF to Telephonics TDH 39 headphones, with MX41AR cushions worn by the subject. The subject's speech was fed back binaurally through the headphones at an intensity of 65 dB above speech reception threshold with a delay of 0.2 second. The intensity level of 65 dB was selected to obtain maximal disruption of speech without undue discomfort to the subjects and was in keeping with levels reported in the literature (Kramer, 1972). The VU meter on the audiometer was used to maintain the signal
at a constant peak level. By setting the volume control so that an average peak reading of 0 dB registered on the VU meter, the intensity of the subject's voice fed back through the headphones could be maintained at 65 dB above his speech reception threshold.

Subjects were assigned randomly to one of two groups, hereafter referred to as the instructed group (IN) and the noninstructed group (NIN). Each group consisted of 20 subjects.

Each subject in the IN group was seated in the subject's room of a sound proofed audiometric suite. Visual and auditory communication between the subject and the investigator was maintained through a window that joins the subject's room and the investigator's room of the suite and through microphones. The subject was instructed to read a reading passage (The Rainbow Passage) silently. After he had read the passage, he was given the opportunity to ask questions concerning pronunciation and vocabulary. The subject was then told to read the passage aloud. This and all subsequent readings were done under DAF conditions. When the first reading was finished, the investigator told the subject the nature of the errors made (nonfluencies or articulation errors or both). A description of the type of nonfluency or articulation error was given to the subject. The subject was given instructions on how to correct the errors during the next reading. The instructions were designed to call the subject's attention to the necessary articulatory movement, position, and contact and to fluency (ongoing, smooth, and effortless linking of speech sounds). The subject was then told to read the passage aloud again. These procedures were used for each successive reading until a total of five readings had been completed. The information given to the subject constituted KR. The same format was used by the investigator to provide information and instruction to all subjects; the investigator read aloud to the subject from a written statement. Actual examples given to the subject depended on the subject's particular errors.

Procedures for the NIN group were the same as for the IN group, with the exception that the subjects in the NIN group were not given KR.

Adaptation scores, based on the decrease in total errors between the first reading and the fifth reading were computed for all subjects in both groups through an analysis of tape recordings made under the DAF conditions. The tape recordings were coded and then scored randomly. A random sample of 20 previously scored tapes was later selected for rescoring in order to assess the reliability of the investigator's judgments. A Spearman rho of 0.95 was obtained between the original analysis and the rescoring. An analysis of each group's performance on the first reading revealed little difference between the groups (the mean number of errors for the NIN group was 5.25 with a standard deviation of 4.07 and the mean number of errors for the IN group was 5.05 with a standard deviation of 3.72).
Equipment and Procedures for Phase II of the Study—Oral Form Discrimination

The stimuli used to measure oral form discrimination consisted of ten plastic geometric forms developed by the National Institute of Dental Research for assessment of oral feedback. The ten forms (Figure 1) are categorized in terms of geometric configurations: triangular (1,2), rectangular (3,4,5), oval (6,7,8), and biconcave (9,10). Pairing forms of the same configuration established “within-class” pairs (e.g., forms 1 and 2); pairing forms of different configurations established “between-class” pairs (e.g., forms 4 and 9).

Kramer (1972) found that using 25 pairs rather than the possible total of 55 pairs accomplished two goals: (1) shortening the time required for testing, thereby reducing possible subject fatigue, which potentially can contribute to erroneous judgments during the test, and (2) creating a test that uses a majority of within-class judgments, making the test more sensitive to differences in oral form discrimination ability. Therefore, 25 pairs were used in the present study.

![Figure 1. Two-dimensional representation of forms used for oral form discrimination testing.](image-url)
Ten of the pairs consisted of identical shapes and were of the within-class type; that is, each of the ten forms was paired with itself. The remaining 15 pairs were established by first pairing each form with every other form, thereby creating 45 pairs and then selecting the eight pairs that constituted the remaining within-class types. Seven pairs of between-class types were selected randomly. Thus, all of the possible within class pairs (18 in number) and seven of the between class pairs constituted the oral form discrimination task (Kramer, 1972).

Each subject was presented with 25 form pairs. The order of presenting the pairs was randomized for each subject. The subject was seated across a table from the investigator in a quiet room. A blindfold was placed over the subject's eyes, thereby preventing the subject from seeing the stimulus forms. Next, the subject was instructed to close his mouth and manipulate the form in his mouth for 5 seconds. The amount of time that each form was in the mouth was measured by the use of a stopwatch. After 5 seconds had elapsed, the subject was told to open his mouth so that the form could

![DAF/KR CONDITIONS](image)

**Figure 2.** Categories of subjects grouped by DAF/KR conditions and OFD rank.
be removed and the second form of the pair placed in his mouth. The subject was told to manipulate the second form in his mouth for 5 seconds, after which the form was removed. Immediately after the second form was removed, the subject was asked if the two forms were the same or different. This procedure was followed for each of the 25 form pairs. A point was awarded for each correct identification; a perfect score was 25.

For the present study, each of the 40 subjects was ranked in oral form discrimination ability. Subjects whose scores were at least one standard deviation above the mean OFD score were ranked as high; subjects whose scores were between plus one and minus one standard deviation from the mean were ranked as average; subjects whose scores were one or more standard deviations below the mean were ranked as low. The mean OFD score for the subjects ranked high was 23.7; for the subjects ranked average, the mean was 20.5; and for the subjects ranked low, the mean was 16.8. A series of t-tests indicated that the groups differed from one another on the basis of OFD scores. Further, the mean scores for the high and low groups were similar to Kramer's (1972) high and low group mean scores of 23 and 17.6, respectively (computed from raw data reported by Kramer).

The mean OFD score for the entire sample was 20.5, the median score was 20.5, the range was 10, and the standard deviation was 2.24 (slightly less than one-fourth of the range). These figures indicate a generally symmetrical distribution and a sample whose nature is representative of what is found in most statistical work (Ventry and Schiavetti, 1980:136).

The subjects were divided into the six categories described in Figure 2.

RESULTS
A factorial design for two-factors analysis of variance was performed using adaptation to DAF as the dependent variable. Results indicated that although the probability that conditions of instruction (KR) had an effect on adaptation was not statistically significant, the probability that the OFD groups differed in adaptation to DAF and the probability of an interaction between OFD scores and instruction (KR) were statistically significant, the respective values being $F = 0.00065, df = 1/34, \text{NS}; F = 4.69, df = 2/34, p < .025; F = 7.04, df = 2/34, p < .005$. A separate analysis of variance indicated that both the IN and the NIN groups adapted to DAF over the five readings.

Table 1 indicates the percentages of adaptation to DAF for subjects ranked high, average, and low in oral form discrimination (OFD). Several trends can be seen from an analysis of the descriptive data in the table: Subjects in the group that received instruction (KR), the IN group, had a higher percentage of adaptation than did subjects in the noninstructed
group (NIN); in both the IN and the NIN groups, adaptation percentages decreased as OFD ranking decreased; within OFD ranks, subjects who received instruction (KR) had higher percentages of adaptation than did the noninstructed subjects; and across the IN and NIN groups, NIN subjects with high and average OFD ranking had higher percentages of adaptation than did IN subjects with a low OFD ranking.

Given the findings of a statistically significant probability of a difference among OFD rankings and of an interaction between OFD and instruction (KR), a post-hoc analysis of the data was made using the Newman–Keuls Multiple Range Test. The intent was to determine which OFD ranks across each KR condition (NIN/IN) differed from one another statistically in percentage of adaptation. The results are displayed in Table 2. Of the 15 pairs of means, five proved to differ from one another at probability levels that were statistically significant.

Table 2. Newman–Keuls' Critical Differences Needed for Significance and Differences between Adaptation Means

<table>
<thead>
<tr>
<th>Group</th>
<th>Critical difference</th>
<th>$\bar{x}_1 - \bar{x}_2$</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low NIN vs. High IN</td>
<td>80.93</td>
<td>-13.33–88.88</td>
<td>102.21*</td>
</tr>
<tr>
<td>Low NIN vs. Average IN</td>
<td>77.16</td>
<td>-13.33–71.79</td>
<td>85.12*</td>
</tr>
<tr>
<td>Low NIN vs. Low IN</td>
<td>54.39</td>
<td>-13.33–47.22</td>
<td>60.55*</td>
</tr>
<tr>
<td>Average NIN vs. Low IN</td>
<td>54.39</td>
<td>47.22–58.88</td>
<td>11.66</td>
</tr>
<tr>
<td>Average NIN vs. Average IN</td>
<td>65.68</td>
<td>58.88–71.79</td>
<td>12.91</td>
</tr>
<tr>
<td>Average NIN vs. High IN</td>
<td>72.46</td>
<td>58.88–88.88</td>
<td>30.00</td>
</tr>
<tr>
<td>High NIN vs. Low IN</td>
<td>65.68</td>
<td>47.22–70.83</td>
<td>23.61</td>
</tr>
<tr>
<td>High NIN vs. Average IN</td>
<td>54.39</td>
<td>70.83–71.79</td>
<td>.96</td>
</tr>
<tr>
<td>High NIN vs. High IN</td>
<td>65.68</td>
<td>70.83–88.88</td>
<td>18.05</td>
</tr>
<tr>
<td>Low NIN vs. Average NIN</td>
<td>65.68</td>
<td>-13.33–58.88</td>
<td>72.21*</td>
</tr>
<tr>
<td>Low NIN vs. High NIN</td>
<td>72.46</td>
<td>-13.33–70.83</td>
<td>84.16*</td>
</tr>
<tr>
<td>Average NIN vs. High NIN</td>
<td>54.39</td>
<td>58.88–70.83</td>
<td>11.95</td>
</tr>
<tr>
<td>Low IN vs. Average IN</td>
<td>72.46</td>
<td>47.22–71.79</td>
<td>24.57</td>
</tr>
<tr>
<td>Low IN vs. High IN</td>
<td>77.16</td>
<td>47.22–88.88</td>
<td>41.66</td>
</tr>
<tr>
<td>Average IN vs. High IN</td>
<td>54.39</td>
<td>71.79–88.88</td>
<td>17.09</td>
</tr>
</tbody>
</table>

* $p < .05$
An examination of Table 2 indicates that the IN groups with low, average, and high OFD rankings achieved higher levels of adaptation than did the NIN group with a low OFD ranking, but that there were no statistical differences in adaptation between any of the IN groups and the NIN groups with average and high OFD rankings. Furthermore, subjects in the NIN group with low adaptation scores did not adapt as much as did subjects in the NIN groups with average to high OFD scores. It would appear that not receiving instruction (KR) coupled with low OFD scores defines those subjects in the present study who did most poorly under DAF.

CONCLUSIONS

The purely descriptive and the statistical analyses in the present study suggest that oral feedback plays some part in a speaker's ability to adapt to the disruptive effects of DAF and that explanations of the adaptation effect need to consider oral as well as auditory monitoring of speech production. Optimal adjustment to DAF appeared to occur in subjects who were aware of intrinsic oral cues or who were helped to become aware of those cues, but did not occur in subjects who had low levels of awareness and at the same time received no help in becoming aware of cues.

The findings of the present study appear to echo Spilka's (1954) theoretical hypothesis that

... the manner and degrees in which speech and voice are altered by delayed sidetone are a function of (1) emphasis on either exteroceptive (auditory) cues or the proprioceptive (kinesthetic) cues of speech, and (2) the ability to shift from one set of these cues to the other (p. 492).

Spilka's hypothesis offers a framework within which to consider the possible interaction between auditory and proprioceptive–kinesthetic feedback:

In developing an explanation of the apparently great individual variation in ability to attend to kinesthetic (proprioceptive) speech cues, the position is taken that all adjustment is, in part, a function of a balance between attention to external and internal cues and the ability to vary such attention as conditions change (Spilka, 1954:492).

The research described herein is based on part of a broader study done for the author's doctoral dissertation at New York University under the direction of Dr. John P. Burke. Other results of that study may be found in Attanasio (1977; 1981).
REFERENCES


