Repetition of Self-Generated Utterances in Conduction Aphasia

Katharine H. Odell
Jill Bonkoski
Jennifer Mello
University of Wisconsin–Madison

Verbal repetition performance was analyzed in four individuals with conduction aphasia. Comparisons at the sentence level included novel versus self-generated, conceptually chained utterances, and spontaneous, narrative utterances versus their repetition. No significant differences were found; however, repetition of self-generated utterances was superior to novel utterances for most subjects. For three speakers, repetition of their own utterances did not show decline relative to spontaneous speech. Error analysis of original and repeated productions showed that speakers produced more instances of preliminary speech errors and word and sound level errors in repetition. Discussion includes comments on assessment and treatment implications.

Aphasologists concur that conduction aphasia (CA) is characterized partly by verbal repetition that is notably inferior to spontaneous speech (Goodglass & Kaplan, 1983; Green & Howes, 1977). Primary types of errors in CA repetition include literal paraphasias and multiple, sequential verbal attempts to produce words or phrases (Goodglass, 1992). The repetition deficit is documented by comparing performance in spontaneous speech and repetition of multisyllable words and sentences, such as the low and high probability sentences found in the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983). However, the degree and consistency of repetition deficit and the nature of differences between spontaneous and repeated speech at phrase and sentence levels have not been well described.

Utterances with presumed strong internal representations, such as those the subject has himself recently created and produced, may pose less difficulty in repetition than novel utterances (Joanette, Keller, & Lecours, 1980). In repetition of self-generated utterances, speakers presumably have advantages of heightened meaning as well as priming of phonological and motoric elements of speech output. Indirect support for this notion comes from Gardner & Winner's (1978) finding that individuals with CA were particularly sensitive to the positive influence of meaningfulness of items to be repeated; items with greater meaning, such as nouns, were significantly easier to repeat than those with less meaning, such as nonsense syllables. One might speculate that the meaningfulness of items to be repeated is increased in a situation where propositional units have been created in the language processing system of the speaker himself. Another aspect of the meaningfulness component that may facilitate repetition efforts is the conceptual relationship among utterances. Speech efforts that are linked in meaning, as in a narrative task, may further solidify the interrelated access of meaning, phonological form, and syntactic form, resulting in greater coherency and fluency in output. Therefore, analysis of the repetition integrity of subjects' own spontaneous utterances that are conceptually chained may contribute to understanding this deficit.

This study analyzed various features in the speech of four subjects diagnosed with CA. One general goal of the study was to compare performance in repetition of novel utterances and of the speaker's self-generated utterances. A second goal was to compare quantitatively and qualitatively the speech in spontaneous picture description and in repetition of the same utterances. Performance in the two conditions was analyzed to determine overall severity of production and the nature of errors.

Method
Subjects

All participants were native speakers of English, premedially right-handed (by self-report), and at least high school graduates. All were mildly aphasic, as indicated by their rather high scores on the Porch Index of Communicative Ability (PICA; Porch, 1967), the Revised Token Test (RTT; McNeil & Prescott, 1978), and the auditory comprehension subtests on the BDAE (Goodglass & Kaplan, 1983). There was no evidence of dysarthria, as revealed by a structural-functional exam (Rosenbek & Wertz, 1976), nor of apraxia of speech, as indicated by performance on the Apraxia Battery for Adults (Dabul, 1979). Melody of speech was largely preserved, as indicated by speech ratings ranging from 5 to 7 on the BDAE (Goodglass & Kaplan, 1983). Importantly, all participants repeated accurately three or fewer of the 8 high and 8 low probability utterances in the "Repeated Phrases" subtest of the BDAE. These traits, especially the perceived discrepancy between poor repetition and better spontaneous speech, are consistent with the typical diagnostic picture for CA. Biographical and test performance data are detailed in Table 1.

Procedure

Performance on the Repeating Phrases subtest of the BDAE (Goodglass & Kaplan, 1983) was used for this study. Participants spontaneously described the Cookie Theft picture from the BDAE (Goodglass & Kaplan, 1983). After a short delay (10-20 minutes), each person repeated his own original utterances, which were presented orally to the speakers in the same phrasal units in which they were initially produced. One model was provided. Speakers were not constrained by time in their repetitions. No feedback was given to the speakers. Verbal reformulations and linguistic or phonetic errors in original utterances were not presented for repetition. Productions were audiorecorded and later narroly transcribed by two trained transcribers. Error analysis was conducted by consensus among three judges.

Reliability

Perceptual judgments of preliminary attempts and narrow phonetic transcriptions of sound segments, relying on the principles...
of the International Phonetic Association (1949), of all original and two repetition Cookie Theft descriptions were executed by two transcribers with proven high intertranscriber reliability (Odell, McNeil, Rosenbek, & Hunter, 1991). Consensus transcriptions for each segment were developed for each sound segment, following guidelines in Shriberg, Kwiatkowski, and Hoffman (1984). One of the listeners resubscribed a subset of sentences or fragments from the spontaneous and repetition speech samples (approximately 25% of each corpus) of each speaker a second time, at an interval of no less than 6 months after the original transcription (Time 1). The number of times individual items (sound segments, preliminary attempts) agreed on Time 1 and Time 2 was calculated at 83%.

One of these listeners, with proven intrancriber reliability, transcribed the repetition condition speech samples for Speakers 3 and 4. Intrajudge reliability was determined by calculating the percentage of item-to-item agreement on perceived sound segments on Time 1 and Time 2. A point no less than 4 months after Time 1. Item-to-item agreement from Time 1 to Time 2, at the level of narrow phonetic transcription, was calculated as 83%.

Reliability of error analysis was determined by reanalyzing all transcripts a second time, no less than 3 months after the first analysis. Reliability was calculated at 100% for severity and at 91% for error type.

Analyses

Four measures of repetition accuracy were calculated, as listed in Table 2. Any deviation from the model, including lexical, grammatical, phonological, or phonetic variations, designated the repetition effort as inaccurate. Two utterance level analyses were done. First, repetition performance in the Cookie Theft description was compared to repetition performance in the BDAE Repeating Phrases subtest. The low probability sentences on this subtest are surely novel, but the high probability ones are not likely novel, although at the time of the examination subjects had not actually spoken the phrases. Second, spontaneous production and repetition of these same utterances in the Cookie Theft task were compared.

Analyses below the sentence level detailed the magnitude of repetition problems, obscured at the more global level because a single error in one word was sufficient to identify a failure in sentence repetition. The analyses also enabled studies of self-corrective behavior and error types. Two types of word level accuracy analyses were done, involving only the spontaneous and repetition conditions of the Cookie Theft task. One analysis determined the quantity of word errors produced, when the words were not preceded by a preliminary effort at production. Essentially, this analysis sought to discover how well the speakers produced words on the first try. The other word level analysis calculated word accuracy in final productions (final try), that is, in words ultimately accepted by the speaker, even if the words had been preceded by a preliminary attempt and, therefore, constituted kinds of self-correction. Basically, the experimental question was: “To what degree did the speaker ever attain accurate production?”

In addition to these four accuracy measures, the nature of misproductions when errors occurred was examined. Error types are listed in Table 3 and defined in Appendix A. One analysis compared the quantity of preliminary speech efforts in the spontaneous and repetition conditions of the Cookie Theft task. Preliminary efforts may be considered self-corrections, preparatory to final productions. The category of preliminary speech effort includes formulations such as recognizable English sound segments, syllables, and words. These units are often referred to as either “self-corrections” or “approximations”; multiple, sequential productions are “successive approximations” (Joanette et al., 1980). Because the literature offers no principle on which to separate subphonemic sound such as laryngealizations or fillers (“uh”) from more obviously phonemic level productions, this category also encompasses both these types of extraneous sound production.

Another analysis of error type assessed the frequency of various types of errors that have been noted in the CA literature (Artila & Rosselli, 1993; Gardner & Winner, 1978; Goodglass & Kaplan, 1983). Ten types of error evident in final productions (i.e., not within preliminary speech efforts) in spontaneous and repetition conditions were tabulated.

Results

Summary data on accuracy were submitted to nonparametric statistical analyses, because the small number of subjects and the use of proportional data made untenable the assumption of a normal distribution of the data, necessary in the use of parametric statistics. All comparisons involved the Wilcoxon signed ranks test (Siegel & Castellan, 1988). To account for the multiple tests, a conservative alpha of $p < 0.025$ was adopted.

Level of Accuracy

The Wilcoxon signed ranks test indicated a nonsignificant difference between repetition performance in self-generated versus non-self-generated, or novel, utterances. Nevertheless, repetition performance was better in self-generated utterances for all speakers (Table 2). For the group, accurate repetition averaged 25% for self-generated utterances but only 9% for the novel BDAE utterances. Speaker 1
TABLE 2. Accuracy measurements. Three tabulations are listed: percentage of repetition accuracy for the BDAE Repeating Phrases subtest; percentage of sentence and percentage of word level accuracy of Cookie Theft description production for all speakers. Numbers outside of parentheses are percentages of the actual counts encompassed in the parentheses.

<table>
<thead>
<tr>
<th>Unit of analysis</th>
<th>Speakers</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Group Mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S1</td>
<td>S2</td>
<td>S3</td>
<td>S4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDAE Repetition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cookie Theft Spontaneous Production Performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence, sentence fragments</td>
<td>6 (1/16)</td>
<td>6 (1/16)</td>
<td>6 (1/16)</td>
<td>19 (3/16)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Single word, without preliminary efforts (&quot;first try&quot;)</td>
<td>50 (7/14)</td>
<td>0 (0/6)</td>
<td>27 (4/15)</td>
<td>12 (2/17)</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Single word, with preliminary effort (&quot;final try&quot;)</td>
<td>91 (72/79)</td>
<td>56 (38/88)</td>
<td>56 (41/73)</td>
<td>77 (61/79)</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Cookie Theft Repetition (self-generated utterances)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence, sentence fragments</td>
<td>7 (1/14)</td>
<td>29 (27/9)</td>
<td>27 (4/15)</td>
<td>35 (6/17)</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Single words, without preliminary efforts (&quot;first try&quot;)</td>
<td>63 (50/79)</td>
<td>65 (44/88)</td>
<td>58 (42/72)</td>
<td>82 (65/79)</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Single words, with preliminary efforts (&quot;final try&quot;)</td>
<td>73 (58/79)</td>
<td>86 (60/88)</td>
<td>72 (52/72)</td>
<td>80 (60/79)</td>
<td>78</td>
<td></td>
</tr>
</tbody>
</table>

*The number of sentences presented for repetition to this speaker was less than the original number he produced, due to examiner error.

exhibited the least change between conditions: from 6% in novel to 7% in self-generated utterances. The other speakers were more accurate on self-generated utterances than on novel stimuli by margins of 15 to 20 percentage points (e.g., 6% to 27%).

A separate Wilcoxon signed ranks test revealed no significant difference in sentence level accuracy between the repetition and the spontaneous conditions in the Cookie Theft task. Mean accuracy in the repetition condition (25%) slightly exceeded that in the spontaneous condition (22%). It is noteworthy that two of the four speakers differed from the group performance in the relationship between the two Cookie Theft conditions (Table 2). Speaker 1 declined considerably in the repetition condition, and Speaker 3 performed similarly in both conditions. Speakers 2 and 4 actually improved in repetition, reflecting greater accuracy on a small number (2 or 4) of the 8–15 sentences. Most of the utterances that contributed to the improvement were judged as propositional; that is, improvement reflected more than better repetition of automatic phrases, such as “I think so.”

To examine the influence of multiple attempts, comparisons were made between word level accuracy on first try productions (words not preceded by a preliminary effort) and final try productions (words preceded by a preliminary effort) in both conditions of the Cookie Theft task (Table 2). Wilcoxon signed ranks analyses revealed no significant differences between conditions in either of these comparisons.

The average group improvement in word production from first to final trial was 13 percentage points (from 70% to 83%) in the spontaneous condition and 9 percentage points (from 67% to 78%) in the repetition condition.

**Error Analysis**

Because of the failure to find significant differences in the overall accuracy measures, the summary data on error types were not submitted to statistical analyses but were analyzed descriptively. As a group, the speakers produced slightly more preliminary efforts in repetition (M = 29) than in the spontaneous condition (M = 24). There was considerable intersubject variability, especially in the repetition task. As shown by the individual data in Table 3, Speaker 2 was outstanding in the number of preliminary efforts produced. In addition, the frequency of preliminary speech efforts in the two conditions varied across speakers. Of the three speakers who produced more preliminary efforts in repetition, two (Speakers 2 and 3) generated one and a half times as many in repetition as in spontaneous production. In contrast, Speaker 4 produced about three times as many preliminary efforts in the spontaneous versus the repetition condition.

The vast majority of preliminary efforts were English phonological units (sound segments, syllables, words). For instance, speakers typically produced[dz2...dz2...dz2...dz2...dz2...dz2...dz2...dz2...]rather than nonphonemic level productions such as "uh" and laryngealizations. Preliminary efforts preceded all form classes, but noun and verb targets were more frequently identified than adjectives, prepositions, or pronouns. When closed class words were involved in preliminary efforts, the attempt was equally likely to be composed of the entire word as of a word fragment, and the attempt was repeated fewer than four times. However, the attempts to produce open class words were often manifest as many repetitions of initial sound segments or syllables. Because each instance of a single sound segment (or syllable) repetition was included in the total, open class preliminary attempts dominated in number.

Concerning word errors, as shown in Table 3, the group, and each individual, produced more word and sound level errors in repetition (M = 16) than in spontaneous speech (M = 8). There was considerable intersubject variation, with ranges from 5–19 instances in spontaneous, and from 9–21 in repetition. The tendency for more word and sound level errors in repetition was greatest for Speaker 1, who made almost four times as many errors in repetition, and least for Speakers 2 and 4, whose error rates in the two conditions were about the same.

The frequency of each word and sound level error type was analyzed (Table 3). In the repetition condition, the most common errors for the group were word omissions, sound segment distortions, and literal paraphasias. In the spontaneous condition, the same three items constituted the top three positions; however, the order of primacy differed: sound segment distortions exceeded word omissions, but literal paraphasias again were third in frequency.
TABLE 3. Numerical frequency of subsentence level error types (preliminary efforts, word level, sound segment level) for all speakers in the spontaneous and repetition conditions of the Cookie Theft task.

<table>
<thead>
<tr>
<th>Unit of analysis</th>
<th>Speakers</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>Group Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S1</td>
<td>R</td>
<td>S</td>
<td>R</td>
<td>S</td>
</tr>
<tr>
<td>A. Total Preliminary Efforts</td>
<td>19</td>
<td>21</td>
<td>48</td>
<td>76</td>
<td>8</td>
</tr>
<tr>
<td>B. Word Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>verbal paraphasia</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>literal paraphasia</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>syntactic elaboration</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>syntactic simplification</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>word omission</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>word addition</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>neologism</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>multword substitution</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total Word</td>
<td>2</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>C. Sound Segment Distortions</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Total Word and Sound</td>
<td>5</td>
<td>19</td>
<td>6</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

*Spontaneous condition
*Repetition condition
*Restricted to bound morpheme level errors and contractions
*Restricted to bound morpheme level errors and contractions

Word omissions involved primarily determiners but occasionally content words, most occurring in utterance-initial positions.

Discussion

Sentence Level Accuracy

All subjects did show better repetition of self-generated utterances than novel utterances, although the difference was not statistically significant. Several traits of the Cookie Theft task may have given the advantage to repetition of these types of utterance. In comparison with the BDAE sentences, the Cookie Theft utterances were linguistically less complex, using more frequently used words and less complexity of phonology and articulation. Furthermore, the sentences presumably possessed relatively strong internal cognitive, linguistic, and motoric representations, as they were produced by the speakers themselves only moments before the repetition. Also, they were conceptually chained, as is obligatory in narrative tasks. The design of this experiment, however, does not allow the determination of whether the reduced linguistic complexity, the self-generated aspect of the utterances, the conceptually chained aspect, or some combination of these was influential in improving repetition. In any of these cases, nevertheless, for the group, repetition of self-generated and conceptually chained utterances appeared to buttress the depressed repetition skills often seen in test situations involving novel utterances. Assessments of repetition skill that rely only on novel stimuli may well be valuable for differential diagnosis but may also understate some repetition skills.

Speaker 1 alone did not demonstrate notable improvement in repetition of self-generated utterances. Speaker 1 scored slightly lower on all four language measures, suggesting that the poorer the language control, the less able an individual with conduction aphasia is to benefit from or, perhaps, to generate a strong internal representation in repetition tasks. The fact that the three more mildly aphasic subjects performed better on repetition of self-generated rather than novel utterances has implications for treatment. Clinicians may be able to capitalize on the ability of some patients with conduction aphasia to develop internal representations of their own utterances to assist in repetition, typically a useful tool in aphasia treatment paradigms. However, time postonset of brain damage, potentially affecting the development of new speaking habits, can be discounted as having a notable impact on repetition achievement. A long time postonset was associated with both low and high levels of repetition improvement on self-generated utterances.

Similarly, in the Cookie Thief tasks, for three speakers, accuracy of sentence level repetition was similar or better than their spontaneous efforts; the vigor of repetition can perhaps be attributed again to the presumably heightened strength of language and speech representations generated in the spontaneous condition. However, for Speaker 1, whose performance declined in repetition as expected in conduction aphasia, other accounts must be sought. Less robust language skills may have contributed, but it is also relevant to consider other factors, such as short-term memory deficits, in this case.

Word Level Accuracy

In the Cookie Theft tasks, all speakers demonstrated higher percentages of successful repetition of individual words (within utterances) than of sentences or sentence fragments. Thus, an individual word approach to comparison showed fair command of word repetition, averaging about 2/3 correct (first try). Speakers heard entire utterances as models, so the full representational and memory burden was not minimal. However, if output processing advanced word-by-word, then the units for production were shorter than full utterances, thus perhaps facilitating production.

Repetition of individual words improved as speakers progressed through several efforts to reach final targets, consistent with the familiar conduites d'approche speech phenomenon (Goodglass & Kaplan, 1983; Joanne et al., 1980). The peculiar deficit in repetition was again evident in that the ameliorative effect of successive efforts was less pronounced in Cookie Theft repetition than in spontaneous production.

Error Analyses

Despite the generally equivalent or better repetition of Cookie Theft sentence level utterances for three speakers, there were more instances of word level errors (preliminary efforts and errors) in repetition for three speakers (1, 2, and 3).

Preliminary attempts were associated more often with open class words than with closed class words. Although they were the source of most preliminary efforts, the intended open class lexical units remained stable, as evidenced by the low incidence of verbal paraphasias within preliminary efforts or in final productions. Attempts were mainly composed of individual sound segments or short sequences of word fragments that preserved a substantial number of the sound features of the target words. It seems that the phonological aspects of the targets were feebly, or perhaps initially erroneously, activated. Thus, in terms of a model
of word production, the data from these speakers are consistent with the claim that, in conduction aphasia, the stage of lexical semantics functions more adequately than a postlexical stage (potentially either the stage at which lexical phonology is retrieved from the lexical semantic units (Garrett, 1984) or the more peripheral stage of phonemic planning (Cigan, Varner, & Baker, 1986; Kohn, 1992). With these data, it is difficult to address meaningfully the issue of whether conduction aphasia reflects impairment at the stage of phonological retrieval or of phonemic planning, but differentiation of these levels by evidence from speech error data is an important future endeavor, theoretically and perhaps clinically.

In contrast to expectations derived from the literature, literal paraphasias did not predominate as an error type in any subject in either condition. Several investigators have suggested that two subtypes of conduction aphasia exist (e.g., Kohn, 1992; Shallice & Warrington, 1977). Reproduction conduction aphasia presents with many literal paraphasias and relatively good repetition, a symptom profile that has been attributed to deficient speech programming; repetition conduction aphasia, in contrast, presents with few paraphasias and relatively poor repetition, ascribed to reduced short-term memory. Two participants in this study demonstrated interesting differences in repetition performance and paraphasias. Speakers 3 and 4 produced relatively many literal paraphasias, with no substantial sentence-level repetition (Cookie Theft) deficit, whereas Speaker 1 produced only few literal paraphasias and substantial repetition deficit. However, it is difficult to make a strong argument for types based on 4 speakers, producing a small number of utterances; the oft-noted excessive variation across speakers with brain damage may well have produced the differences described here.

The pattern of performance differed among subjects and is not readily explained. In the Cookie Theft repetition, Speakers 3 and 4 produced more word and sound level errors than they did in preliminary speech errors; the opposite trend was apparent in Speaker 2.

Conclusion

Results of this study suggest that, for some speakers with CA, the magnitude of repetition deficit can be reduced by one or more of several factors, including a strong internal representation of the message and a message that is composed of utterances that are related in meaning. For three of the four speakers in this study, overall sentence level accuracy improved, although not significantly, in repetition of self-generated utterances versus novel utterances. In the Cookie Theft tasks, repetition by three of the four speakers did not decline relative to the spontaneous condition, as expected in CA.

However, the frequency of word level errors (especially sound segment distortions, omissions of determiner-type words, and literal paraphasias) and preliminary efforts was greater in the repetition than the spontaneous condition in the Cookie Theft tasks. Thus, reproduction of self-generated utterances did not eliminate the increased production difficulty in repetition.

References


Key Words: conduction aphasia, verbal repetition, language pathology

Appendix A

Definition of Error Types That Were Identified and Tallied

1. Preliminary effort: any verbal unit (reformulations involving several words, subphonemic and phonemic speech productions, laryngealizations, filler such as oh) that is followed by another effort or by the utterance finally produced by the speaker.

2. Literal paraphasia: real or nonsense word, close in sound to target and recognizable as an attempt at the target but involving a sound segment substitution.

3. Verbal paraphasia: real word substitution, close in meaning to target.

4. Multiword substitution: substitution of a phrase for a word.

5. Elaboration: morphosyntactic addition (e.g., "looking"/"look").

6. Simplification: morphosyntactic reduction (e.g., "look"/"looking").

7. Omission: whole word deletion.

8. Addition: whole word addition (e.g., "boy is reaching for the jar"/"boy is reaching").

9. Neologism: nonsense word for which the target could not be determined.

10. Sound segment distortion: segment level deviations from standard phonetic production (not attributable to casual speech).