Background and Purpose—This article is a comprehensive review of aphasia treatment studies for the purpose of investigating the relationship between time postonset of aphasia and response to treatment for aphasia in chronic patients at ≥1 year after symptom onset.

Methods—Studies that demonstrated treatment response (defined as a measurable change in task performance compared with a control task performance) through the use of single-subject design methodologies on measures of verbal output or auditory comprehension were selected. Individual subject data were extracted from the 23 studies that met criteria identifying the subjects as those who received direct continuous therapy for spoken language deficits and whose changes in response to therapy were measurable. Percent of maximum possible change was used as a measurement of outcome.

Results—Nonparametric correlation statistics (Spearman $\rho$) and comparisons of group means (Kruskal–Wallis) were used to compare the relationship between time postonset and improvement. Time postonset at which treatment was initiated did not correlate with response to treatment. No significant differences in response to treatment were found between groups of patients according to times postonset.

Conclusions—Time postonset is not related to response to treatment for aphasia in patients >1 year postonset of aphasia.

(A Stroke. 2006;37:3043-3051.)

Key Words: aphasia ■ rehabilitation ■ speech disorders ■ speech therapy
ypo made a “complete recovery” after 5 months of therapy. Although other cases were not as dramatic, major gains were achieved by language therapy. Progress was even documented for an individual who was 11 ypo.8

Additional evidence also supports the idea that patients who are >12 months poststroke continue to make gains. Based on a meta-analysis of within-effect outcomes of therapy trials begun at various times postonset, Robey8 found that in the acute phase (0 to 3 months), treated patients made twice as much change as untreated subjects. In the “postacute” stage (3 to 12 months), treated patients made 1.5× more improvement than untreated counterparts. In the chronic phase (>1 year), treated patients made 12× as much improvement as untreated counterparts. The amount of change among the chronic patients was relatively modest. In fact, the treated patients in the chronic phase improved as much as the untreated group in the acute phase. Nevertheless, in the chronic phase, the treated patients changed much more than the untreated patients.

Longitudinal studies of language recovery in aphasia provide substantial evidence that some patients with aphasia continue to improve well beyond 1 year. Fitzpatrick,10 for example, conducted a retrospective analysis of confrontation naming and naming in discourse among patients. She examined data from 20 individuals who were seen multiple times after 1 ypo up to as long as 12 ypo. Fitzpatrick found that confrontation naming scores increased over time in subjects who were evaluated at least twice between 1 and 12 ypo. Aphasia type and ypo were significant predictors of confrontation naming scores, but the patients with nonfluent aphasia and those with fluent aphasia did not change at different rates. Results from this study demonstrated that naming abilities improved for as long as 12 ypo in patients with chronic aphasia.

Additional longitudinal data from Hanson et al5 suggest that patients with aphasia may show long-term improvements. Thirty-five patients were divided into groups based on their initial score on the Porch Index of Communicative Abilities, a standardized test of aphasia. The Porch Index of Communicative Abilities was administered to the patients 6 times between 3 and 55 months postonset. All groups of patients continued to improve significantly on speaking tasks when tested at 6, 12, and 24 months postonset, regardless of severity. The group with the most severe impairment in speaking at each test interval made the greatest amount of change, although they remained the most impaired.

Based on these studies, it is clear that patients can show improvement in language ability long after the generally accepted period of spontaneous physiological recovery. However, little is known about how patients respond to individual treatments conducted later and if the amount of change seen later can be related to previous therapy. Although the latter question may be pertinent, the question addressed here assumes that therapy we provide is effective and well suited for an individual patient,9 regardless of previous therapy. We also consider that an adequate control and test task can offset the effects of previous therapy. So, the issue of previous therapy aside, can clinicians expect a certain amount of recovery based on the time postonset at which therapy is conducted? In other words, does the amount of change we see vary based on time postonset in chronic patients?

These questions remain unanswered because there are few available large-scale longitudinal studies that follow patients’ responses to therapies. The ideal method to investigate the overall effects of time postonset on response to treatment would be a mixed-effect, repeated-measures design, following treated patients of a certain aphasia type at repeated re-evaluations at set intervals for years after their stroke, comparing their progress with untreated controls. Because such ideal studies do not exist, this review is an indirect way of synthesizing data gathered from many studies in order to better understand the influence of time postonset on improvement in aphasia. Similar reviews have likewise attempted to isolate the effects of 1 variable’s effect on treatment outcome.11

Purpose

The purpose of this review is to address the following question: Is there a relationship between time postonset at the initiation of therapy and response to treatment in the chronic phase (ie, >1 ypo)? By extracting individual data from treatment studies, we can begin to investigate any patterns that emerge across studies in the amount of response individuals make when treated after 1 ypo. This review can contribute information relating to the question of whether treatment after 1 ypo is warranted and can help to identify the methodological difficulties that arise in attempting to answer this question.

Methods

Data Collection

An extensive review of the literature was conducted to retrieve relevant treatment studies. A literature search was performed using Medline, PsychInfo, ERIC, and the Cochrane Database of Systematic Reviews to identify studies containing the following search terms: “aphasia,” “follow-up studies,” “language,” “recovery,” and “treatment.” Years of publication were restricted to 1985 to 2003. The year 1985 was selected as a cutoff year because more rigorous methodological standards were introduced into the speech–language pathology field during the mid-1980s. The VA Field Advisory Council Evidence-Based Treatment Outcomes12 was also reviewed. These searches yielded ≈100 articles.

Inclusionary and Exclusionary Criteria

The studies included in this review presented individual data for subjects with aphasia who received treatment for disorders of verbal expression or auditory comprehension and who were >1 ypo at the time they were studied. Furthermore, each study in the review was published between 1985 and 2003 and included sufficient detail on treatment outcomes that measurements of change associated with treatment could be calculated for individual subjects. Specific criteria for inclusion/exclusion follow:

1. To attribute change to response to treatment, each study had to demonstrate adequate control. Most often, this
was accomplished through establishing baseline stability of the tested behavior. Many researchers also used multiple baselines showing untreated control tasks.

2. Only those studies that gave specific time postonset data and examined treatment of a spoken language disorder by a speech–language pathologist in a face-to-face situation were reviewed.

3. Only those studies that presented individual pretherapy and post-therapy scores that referred to change in verbal output or auditory comprehension were examined. Studies that addressed therapy for written language disorders (reading and writing) and speech disorders (dysarthria or apraxia of speech) were excluded.

4. Those that examined treatment for psychological, functional, and emotional impact were excluded.

5. Studies that targeted caregivers’ communication abilities were also excluded.

6. All group therapy studies were excluded.

7. Studies were also excluded if they did not include maximum scores for outcome measure or if they could not be derived.

8. Finally, 2 studies were excluded because the journals in which they were published could not be located.

The authors’ consideration of time postonset as a factor in recovery was not a prerequisite for inclusion. However, only studies from which time postonset data for an individual were available were included. Research articles including any aphasia type were included because limiting the review to a specific aphasia type was considered premature, yielding too few articles to compile an adequate review. In addition, controlling for severity was impossible because of the varying testing protocols.

From the original pool of articles, 23 studies fulfilled these criteria, yielding individual data for 57 subjects. Of these studies, 20 investigated treatment for disorders of verbal expression and 3 investigated treatment for disorders of auditory comprehension. The individual data extracted from these 23 studies are presented in Tables 1 and 2.

**Outcome Measure**

For the purposes of this review, the main outcome measure of change in response to treatment was defined as percent of maximum possible change (%MPC). This measure was calculated for each individual based on the reported or calculated raw scores. When accuracy was presented as a percent correct score, information pertaining to the testing instrument was used to calculate the raw scores. When data were presented in graphical form only, individual data were extracted as the visual estimate of data point values. When accuracy was presented as a percent correct score, information pertaining to the testing instrument was used to calculate the raw scores. When data were presented in graphical form only, individual data were extracted as the visual estimate of data point values.

Because of the many ways that scores were reported, several decisions were made about how to calculate pretherapy and post-therapy scores. The pretherapy score (always taken at a point >1 ypo) was defined as the mean of the baseline scores or the reported pretherapy score (if only 1 was given). The final outcome measure was defined as the last measurement taken before a maintenance phase (when given). The post-test score was not derived from a mean of the assessments made during the therapy phase because the mean score would not adequately account for the duration of treatment and the cumulative effect, which would theoretically be evident from the final score. For some subjects, their final scores may not reflect their overall gains or abilities. However, the large sample size of this review (n=57) may compensate for any effect of using the final measurement rather than the mean of measurements taken during therapy.

The %MPC measurement was chosen as the primary indicator of treatment response for a number of reasons. Simple percent change (as opposed to %MPC) is not a sufficient measure of change on its own because it is affected by the interaction between the number of items on a test and a subject’s original score. For example, a subject who improves from a score of 20 to a score of 40 of a maximum of 100 points increases 100% over the baseline score. However, a person who goes from a score of 40 to a score of 60 of 100 points only increases 50% over baseline. Thus, percent change penalizes the subject who originally performs higher. Using a different measure of change (%MPC) prevents the “penalty” for higher initial scores as demonstrated in the above example. In fact, the closer a person is to a maximum score, the more difficult it becomes for him/her to show improvement, so an absolute change of X points is not always equivalent. Thus, a person who starts higher and makes the same absolute gain as a person whose initial score is lower should actually receive a higher designation because he/she has less opportunity for improvement (ie, fewer test items on which to perform well). Thus, %MPC best reflects assumptions about test performance that would be unaccounted for by using percent change. There is no way to absolutely represent change given the varied outcome measurements reported in the studies selected. %MPC actually penalizes subjects for low initial performance. For example, if a patient starts with a low score and doubles his/her score, this may be clinically significant. However, if he/she made small gains relative to the maximum possible, the %MPC will be very low. To account for the conflict between the cost and benefit of using %MPC, ideally, it would be best to control for initial severity by examining subjects with similar initial scores. However, the limited availability of studies for which this is possible would reduce the sample size and minimize any statistical power. It was decided that because the inten-
TABLE 1. Verbal Output Treatment Studies

<table>
<thead>
<tr>
<th>Study Authors</th>
<th>Subject No.</th>
<th>Aphasia Type</th>
<th>Months Postonset</th>
<th>%MPC</th>
<th>Outcome Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annoni et al14</td>
<td>1</td>
<td>Mixed</td>
<td>36</td>
<td>47%</td>
<td>Boston</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Nonfluent</td>
<td>84</td>
<td>69%</td>
<td>Naming Test</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Wernicke's</td>
<td>36</td>
<td>36%</td>
<td>Broca's</td>
</tr>
<tr>
<td>Best et al15</td>
<td>4</td>
<td>Anomic</td>
<td>26</td>
<td>47%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Boyle and Coelho16</td>
<td>5</td>
<td>Broca's</td>
<td>65</td>
<td>89%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Conley and Coelho17</td>
<td>6</td>
<td>Broca's</td>
<td>96</td>
<td>67%</td>
<td>Study specific</td>
</tr>
<tr>
<td>DeDe et al18</td>
<td>7</td>
<td>Nonfluent</td>
<td>48</td>
<td>22%</td>
<td>PAL verbal naming</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Broca's</td>
<td>88</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Broca's</td>
<td>177</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>Doyle et al19</td>
<td>10</td>
<td>Broca's</td>
<td>175</td>
<td>5%</td>
<td>Northwestern Syntax</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Broca’s</td>
<td>30</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Broca’s</td>
<td>79</td>
<td>24%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Broca’s</td>
<td>36</td>
<td>76%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Broca’s</td>
<td>66</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>Drew and Thompson20</td>
<td>15</td>
<td>Broca’s</td>
<td>30</td>
<td>64%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Hesketh21</td>
<td>16</td>
<td>Broca’s</td>
<td>84</td>
<td>86%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Hickin et al22</td>
<td>17</td>
<td>Broca’s</td>
<td>72</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Anomic</td>
<td>36</td>
<td>87%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>Mixed/Wernicke’s</td>
<td>60</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Anomic</td>
<td>60</td>
<td>58%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>Anomic</td>
<td>24</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>Anomic</td>
<td>36</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>Broca’s</td>
<td>36</td>
<td>14%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>Broca’s</td>
<td>96</td>
<td>12%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Kearns23</td>
<td>25</td>
<td>Broca’s</td>
<td>36</td>
<td>19%</td>
<td>PICA</td>
</tr>
<tr>
<td>Li et al24</td>
<td>26</td>
<td>Conduction</td>
<td>24</td>
<td>30%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Marshall et al25</td>
<td>27</td>
<td>Agrammatic</td>
<td>168</td>
<td>48%</td>
<td>Study specific</td>
</tr>
<tr>
<td>McNeil et al26</td>
<td>28</td>
<td>Anomic</td>
<td>228</td>
<td>88%</td>
<td>Study specific</td>
</tr>
<tr>
<td>McNeil et al27</td>
<td>29</td>
<td>Anomic</td>
<td>36</td>
<td>92%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Nickels and Best28</td>
<td>30</td>
<td>Nonfluent</td>
<td>48</td>
<td>47%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>Fluent</td>
<td>48</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>Fluent</td>
<td>72</td>
<td>40%</td>
<td>Study specific</td>
</tr>
<tr>
<td>Roberts and Wertz29</td>
<td>33</td>
<td>Transcort. Motor</td>
<td>185</td>
<td>22%</td>
<td>WAB</td>
</tr>
<tr>
<td>Robson et al30</td>
<td>34</td>
<td>Jargon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>Broca’s</td>
<td>43</td>
<td>42%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>Wernicke's</td>
<td>20</td>
<td>31%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>Broca’s</td>
<td>69</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>Broca’s</td>
<td>12</td>
<td>87%</td>
<td></td>
</tr>
<tr>
<td>Springer et al31</td>
<td>39</td>
<td>Broca’s</td>
<td>15</td>
<td>22%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>Broca’s</td>
<td>45</td>
<td>50%</td>
<td>Study specific</td>
</tr>
<tr>
<td></td>
<td>41</td>
<td>Agrammatic</td>
<td>24</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>Agrammatic</td>
<td>28</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>Thompson and McReynolds32</td>
<td>43</td>
<td>Agrammatic</td>
<td>15</td>
<td>36%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>Agrammatic</td>
<td>20</td>
<td>59%</td>
<td>Study specific</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>Broca’s</td>
<td>48</td>
<td>44%</td>
<td></td>
</tr>
<tr>
<td>Wambaugh and Thompson33</td>
<td>46</td>
<td>Broca’s</td>
<td>20</td>
<td>79%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>Broca’s</td>
<td>27</td>
<td>64%</td>
<td>Study specific</td>
</tr>
</tbody>
</table>

PAL indicates psycholinguistic assessment of language; PICA, Porch Index of Communicative Abilities; WAB, western aphasia battery.
tion of this project is not only to describe trends in response to treatment but also to identify issues that affect the analysis of outcome. %MPC was adequate for providing interpretable results despite the few methodological problems discussed here.

**Data Analysis Decisions**

For each subject, the type of therapy received was classified as one that targeted either verbal output or auditory comprehension. When the target of therapy was the semantic system itself, and thus did not clearly target verbal expression or auditory comprehension in isolation, the therapy requirements and the measurement used to assess change were considered to best determine whether to include the study in the group of auditory comprehension studies or verbal output studies.

If several therapies were given and no mention was made of any interruption or break between them, it was assumed that they were continuous. However, if there was a break, only the results of the first therapy were included. For example, Nickels et al.13 gave 2 therapies sequentially for comprehension and production deficits. A 4-week break divided both stages of therapy. To report the gains made by each individual therapy would have represented the subject twice. Therefore, it was decided to report the data pertaining only to the first phase of therapy: the comprehension phase.

Finally, many researchers used the absolute change in performance during therapy tasks as their outcome measure; thus, in many cases, outcome measures were essentially treatment targets as opposed to scores on standardized tests or additional measures of accuracy on control or untreated items. Although this may not be the preferred way to measure treatment effectiveness, excluding these studies would have left few studies for review. Very often, control stimuli (not the control task) were also included to measure generalization. Results for both stimuli types were averaged. By including the control stimuli, the scores become an indicator of overall gain and stasis, which is theoretically more comparable to standardized test scores.

**Results**

**Individual Subject Data**

Tables 1 and 2 provide information on aphasia type, months postonset, and %MPC for each subject from each of the 23 studies included in the review. Table 1 lists the subjects in the studies investigating treatment for verbal output (20 studies), and Table 2 lists those in studies investigating auditory comprehension (3 studies).

**Data Analysis**

Each subject’s %MPC was plotted on scatter plots. Data were represented and analyzed as part of the group of 57 subjects and as part of the smaller treatment-target groups. Individual data were also grouped by ypo to compare statistically response to treatment across various times postonset. To
maintain ≥6 subjects in each group, 6 groups were created. The groups consisted of 7 subjects who were 1 to 2 ypo, 12 who were 2 to 3 ypo, 11 who were 3 to 4 ypo, 11 who were 4 to 6 ypo, 6 who were 6 to 8 ypo, and 10 who were >8 ypo.

A Spearman ρ statistic was used to analyze the relationship between time postonset and %MPC and the Kruskal–Wallis 1-way ANOVA by ranks (KWANOVA) was used to compare performance across groups. The KWANOVA was chosen because of the heterogeneity of the population sampled and represented in this review. Because data were extracted from many sources, no assumption about equal variance could be made. Moreover, the KWANOVA uses ranked data, which is appropriate for this study. Using ranks minimizes the effects of not having item equivalence across the various tests used to measure response to treatment because it reduces the distance between scores on different assessments.

All Studies: Treatment for Verbal Expression and Auditory Comprehension

Figures 1 and 2 show results for all subjects, regardless of the therapy target (ie, auditory comprehension or verbal expression). A Spearman correlation was calculated: \( r_s(56)=0.360 \) (\( P<0.05 \)), indicating no significant relationship between the individual %MPC and time postonset.

The results of the KWANOVA indicated no significant differences among the 6 groups \( H(5, 47)=0.410 \) (\( P>0.05 \)). Note in Figure 2 and Table 3 that all groups made gains of ≥28% of the maximum possible improvement. The large SDs reflect the high degree of variability in scores within each group. The individual data in Figure 1 demonstrate this variability, showing that time postonset bears little to no relationship with response to therapy in this sample. Across the time scale, some patients improved nearly 100% of the maximum possible, whereas 1 declined despite treatment.

One subject who was 19 ypo made gains of 88% of the maximum possible.

Subset of Studies for Treatment of Verbal Expression

Figures 3 and 4 display data for the subset of studies investigating verbal expression, which included 47 subjects. A Spearman ρ was calculated for these data and resulted in \( r_s(46)=0.342 \) (\( P>0.05 \)), indicating no correlation between time postonset and %MPC.

A KWANOVA showed no significant difference between the 6 groups of subjects receiving therapy for verbal expression. \( H(5, 55)=0.698 \) (\( P>0.05 \)). Table 4 reports the means and SDs for each group. Again, the SDs are high, reflecting the variability of the data within each group.

Treatment of Auditory Comprehension

Figure 5 shows individual data from the 3 studies investigating treatment for auditory comprehension.

Neither correlation statistics nor a KWANOVA could be performed on the data from subjects who received therapy for auditory comprehension because there were so few individuals. Despite the small sample size, it appears from Figure 5 that there is wide variability in response among patients who are treated for auditory comprehension and that response does not appear related to time postonset.

Summary of Results

Individual data show wide variability in patient performance that is not related to time postonset. When subjects were grouped by ypo, no significant difference between the groups was found. Therefore, time postonset in patients who are >1 ypo does not correlate with patient response to treatment. There are too few cases of reported treatment response for therapy targeting auditory comprehension to make a definitive conclusion that time postonset is or is not related to time postonset for these subjects. Claims made by others that prognosis for response to treatment is a function of time postonset, at least in the chronic phase (>1 year), are not substantiated by a comprehensive review of the treatment literature.

Analysis of the group means indicates the possibility that response does decline at some point after 8 ypo. This group (>8 ypo) showed the least amount of %MPC (28%) and were ≥10 percentage points below the means of the other groups.
on this measure. An increased sample size would perhaps magnify any possible effects of time postonset in this group. Nevertheless, the presence of high-scoring individuals in this group warrants cautious interpretation of the group performance of patients who are ≥8 years after stroke.

**Discussion**

Justification for initiating therapy in the chronic stage is often questioned by clinicians as well as insurers. However, the analysis of data provided by the treatment outcome studies examined in this review found no significant relationship between time postonset and response to treatment. In fact, average changes within the groups when all subjects were included ranged from 28% of maximum possible change to 47%. These results suggest that response to therapy is not correlated with time postonset in patients who are >1 ypo and who are receiving therapy for verbal output deficits. Thus, based on the results of this study, time postonset should not be considered a significant predictor of response to treatment in patients receiving treatment for verbal output disorders who are >1 ypo.

Within a given period of time postonset, subjects appear to make vastly variable degrees of improvement. The plotted individual data reveal this variability, which is also reflected in the large SDs seen in the ypo groups. No significant differences between the ypo groups were found in the analysis of all 23 studies or only in the subset of studies investigating treatment of verbal expression. All else being equal, this result indicates equal potential for improvement in people with chronic aphasia, regardless of the time postonset at which treatment is initiated.

**External Factors and Methodological Considerations**

The interaction between time postonset and other possible factors in recovery, such as aphasia severity at onset and at time of therapy, intensity of therapy, aphasia type, therapy type, and premorbid intelligence cannot be overlooked. To accommodate a level of control that would account for the many confounding variables that intrude in all aphasia outcome studies would have meant greatly limiting the number of subjects studied, thereby reducing the statistical power of the results. In fact, the large number of studies and variability among subjects was considered the most effective way to balance out the sample and eliminate the effects of other confounding variables. Of course, in an ideal study, these variables would have been systematically controlled.

Sarno3 best summarizes the difficulty in controlling for the effects of 1 factor when examining another in the following statement: “Subject selection criteria; subject classification; aphasia severity; characteristics of control groups; premorbid language competence; age, education, socioeconomic status, intellectual level, and personality of patient; time since onset, frequency and intensity of treatment; nature of interventions; and the skill of the therapist are some of the factors which must be considered in aphasia recovery and rehabilitation research. In addition, these variables interact with each other and cannot be studied in isolation.”

To this list, we would also add that previous therapy could have an effect on the results of treatments studied. Another particularly relevant issue in this regard is the file drawer phenomenon, in which researchers do not publish reports of clients who do not make progress. The lack of published data on these subjects may also lead to a misunderstanding of the
effects of time post onset and other possible prognostic indicators of the amount of response to treatment.

We attempted to tease out factors that might potentially skew interpretation of the data by separating auditory comprehension and verbal expression treatment studies and conducting individual analyses of these 2 subsets of studies. However, one might argue that the interaction between type of aphasia and therapy type would still bias the results. For example, the sample that was treated for auditory comprehension included some patients with Broca’s aphasia, and, theoretically, we would expect these patients to score high on baseline tests of auditory comprehension and then perhaps make greater gains. However, note in Table 2 that the scores of the 1 subject with Broca’s aphasia receiving treatment for auditory comprehension declined, resulting in a %MPC of −15. Evidently, the expected interaction between therapy type and aphasia type does not always hold, and trying to control for this interaction would not necessarily be justified. This echoes Robson et al:30 “We are some way from understanding how therapy operates. There appears to be no clear match between the underlying deficit and the form of treatment offered.”

This review also considered the problem of comparing vastly different measures of response to treatment. By giving everyone a score based on %MPC, comparison across tests was facilitated. Additionally, the use of ranked medians in the KWANOVA analysis would theoretically mask the distance between individual scores on different tests. These 2 strategies were deemed sufficient to minimize the effect of comparing different outcome measures.

### Attributing Change to Treatment Response

All studies reviewed for this analysis used experimental designs that attempted to control for natural recovery. It should be noted that some of the studies used an A-B single-subject design, which has received some criticism.36. Other studies used the multiple baseline design, which may better control for spontaneous recovery. This design is based on an assumption that spontaneous recovery is a global phenomenon affecting all tasks. Therefore, stable scores in untreated tasks and improved scores on treated tasks indicate that natural recovery is not uniquely responsible for treatment effects. Although some object to these methodologies and their assumptions, they are well accepted and have been used in much of the treatment efficacy research in aphasia since the 1980s.

The studies presented here purported to look at treatment effects separate from changes related to spontaneous recovery. It is worth restating that comparison with well-matched untreated control subjects across the lifetime would be most informative. However, if we accept the investigators’ controls for the subjects examined here, then we can conclude that treatment was responsible for change. Furthermore, the amount of that change does not appear to be related to the time at which therapy is initiated in patients who are >1 ypo.

### Clinical Significance

Although the results suggest that some chronic patients receiving treatment for verbal output deficits have the potential to improve with treatment, it is not possible to determine from this review how these treatment effects translate to functional communication. Only a handful of the studies reviewed here considered functional measurements of change in addition to measurements of specific linguistic change. Because functional communication measures often assess compensatory modalities as well as natural language functions, it is presumed that functional communication profiles might look quite different from the treatment profiles in this review. Future research might compare measures of functional communication across time with measures of more discrete linguistic tasks.

One can question whether an average of 42% change in the chronic period is clinically significant. At first glance, this amount seems substantial, but the question of clinical significance cannot be adequately answered by this review. In fact, this amount (>40%) may simply reflect the absence of published research studies describing patients who do not make progress. Nevertheless, the high variability among the subjects perhaps counteracts that argument and enhances the generalizability of the reported means.

### Table 4. Mean %MPC and SD by Group for Subjects Treated for Verbal Expression

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Mean %MPC</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–2 ypo</td>
<td>6</td>
<td>52</td>
<td>27</td>
</tr>
<tr>
<td>3–4 ypo</td>
<td>10</td>
<td>50</td>
<td>27</td>
</tr>
<tr>
<td>4–6 ypo</td>
<td>9</td>
<td>49</td>
<td>28</td>
</tr>
<tr>
<td>6–8 ypo</td>
<td>6</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>8–19 ypo</td>
<td>7</td>
<td>35</td>
<td>33</td>
</tr>
</tbody>
</table>
The data indicate a slight (albeit nonsignificant) decrease in mean score in group 6 (8 to 19 ypo) for both the sample of all subjects and the subset of those treated for verbal expression. A larger sample size, and greater number of studies involving patients who are >8 ypo, would elucidate any possible effects of time postonset that may happen in the time after 8 ypo because this group included the widest range of times postonset. However, based on the available literature, it is impossible to claim significant effect of time postonset even in the time period after 8 ypo.

Conclusion
This review of the aphasia treatment literature demonstrates that many people who receive treatment will respond positively to treatment even many ypo of aphasia. Furthermore, there is no correlation between time postonset and amount of change seen with treatment in people with aphasia who are ≥1 year postonset. Therefore, expectations of how well individuals will respond to aphasia treatment, which are based solely on time postonset, are not warranted.

Acknowledgments
We would like to thank Dr Kevin Kearns, PhD, CCC-SLP, for his assistance in the development of this project and Dr James Heaton, PhD, for his input and advice. We are also grateful to Beth Stromstadt, MA, for her statistical expertise.

Disclosures
None.

References
Language Rehabilitation in Chronic Aphasia and Time Postonset: A Review of Single-Subject Data
Aviva Moss and Marjorie Nicholas

Stroke. 2006;37:3043-3051; originally published online November 9, 2006; doi: 10.1161/01.STR.0000249427.74970.15

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2006 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/37/12/3043

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at: http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at: http://stroke.ahajournals.org//subscriptions/