Comparing the outcomes of intensive and non-intensive context-based aphasia treatment

Jacqueline J. Hinckley

University of South Florida, Tampa, FL, USA

Thomas H. Carr

Michigan State University, East Lansing, MI, USA

Background: Intensive rates of treatment have been shown to have positive outcomes but have rarely been directly compared with non-intensive treatment. Certain types of treatment may be more effective at intensive rates than others.

Aims: The purpose of this study was to compare intensive and non-intensive rates of one particular, highly specified type of treatment termed “context-based treatment”.

Methods & Procedures: Thirteen adults with moderately severe aphasia were assigned to either intensive or non-intensive treatment. A battery of assessments was designed to measure the effectiveness of the treatment and the transfer of the treatment to more and less similar contexts.

Outcomes & Results: There was no advantage of intensive treatment for achieving mastery of the trained context, or in transferring those skills to similar environments, or challenging environments.

Conclusions: If replicated, the results could suggest that context-based treatment may be a treatment type of choice when treatment time is limited.

For any aphasia treatment, it is critical to know the appropriate dosage in terms of total amount and rate of treatment. Unfortunately, this is an issue that has received little clinical research attention.

An intensive rate of treatment can be defined as more treatment provided over a shorter amount of time (Hinckley & Craig, 1998). Robey’s (1998) meta-analysis suggested that more treatment was associated with better outcomes, and several empirical studies have demonstrated positive outcomes of intensive rate (e.g. Mackenzie, 1991; Poeck, Huber, & Willmes, 1989; Pulvermuller et al., 2001; Wertz et al., 1986; see Bhogal, Teasell, Foley, & Speechley, 2003, for a recent review). Few studies, however, have been designed to directly compare intensive and non-intensive rates of treatment. In two such studies, positive outcomes of the intensive treatment were reported, but the treatment was not well specified in the published report (Denes, Perazzolo, Piani, & Piccione, 1996; Hinckley & Craig, 1998). Without knowing exactly what activities
comprised the treatment, it is not possible to replicate the treatment or to determine whether the positive outcomes might generalise or are limited to the particular combination of activities used in that study.

The idea that intensive treatment may be more efficient for certain treatment types can be linked to the literature on massed versus distributed training. Massed training refers to training in which there are no breaks between practice items. There are breaks of hours or even days between practice trials in training that is described as distributed (Donovan & Radoshevich, 1999). Meta-analysis of skill training among non-impaired populations suggests that distributed practice is generally better, except when the task to be learned is particularly complex or when the likelihood of errors or forgetting is high (Donovan & Radoshevich, 1999).

Thus, if one assumes that intensive rates correspond more to massed training whereas non-intensive rates correspond more to distributed training, then one should expect that intensive treatment would be effective for some but not all language tasks (more effective for more complex tasks, for example) and with some aphasic populations more than others (more effective when forgetting is a more significant component of the patient’s deficits). This hypothesis from the skill-training literature underscores the need in aphasiology to specify the treatment and investigate various types of treatments at intensive and non-intensive rates. The present study compared intensive and non-intensive rates of a particular treatment type, referred to as context-based treatment, among adults with non-fluent aphasia. Context-based treatment is consistent with the current emphasis in the field on functional outcomes. We selected participants with non-fluent aphasia because their relatively preserved auditory comprehension would facilitate ability to follow task instructions.

METHOD

Design

Participants were assigned to either intensive or non-intensive context-based treatment in a between-groups pre–post test design.

Participants

Participants had a moderate-to-severe non-fluent-type aphasia as determined by interpretation of scores on the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) or the Western Aphasia Battery (Kertesz, 1982) due to a single left hemispheric CVA. They also had sufficient physical and cognitive endurance to participate in an intensive treatment regimen. All participants were monolingual speakers of English who were premorbidly right-handed. None reported a history of other neurologic disease, psychiatric diagnosis, or substance abuse. Participants were at least 3 months post-onset of stroke and aphasia.

A total of 13 adults with aphasia met these criteria and were assigned to either intensive or non-intensive treatment. Of these, 12 were diagnosed with Broca’s aphasia, and one with transcortical motor aphasia. All of the participants were rated between 1 and 3 on the 6-point Severity Rating Scale of the BDAE. Socioeconomic status was estimated

---

1 The second edition of the Boston Diagnostic Aphasia Examination was used because the third edition was not available at the time that the study began.
using the Four Factor Index (Hollingshead, 1975), which is weighted on education and occupation.

The two groups were comparable for age, time post-onset, socioeconomic status, and aphasia severity. The means, standard deviations, and ranges of these descriptors for the two groups are provided in Table 1.

Both groups demonstrated relatively preserved cognitive abilities as evidenced by their performance on tasks associated with the Global Aphasia Neuropsychological Battery (van Mourik, Verschaeve, Boon, Paquier, & van Harskamp, 1992). Scores for all participants in both groups were high for a visual cancellation task (range = 97–100%), and for the object recognition subtest of the Rivermead Behavioral Memory Test (Wilson, Cockburn, & Baddeley, 1985) (range = 71–100%). Performances were more variable across participants on the Ravens Coloured Progressive Matrices (Raven, Court, & Raven, 1979) (range = 41–100%). On the Wisconsin Card Sort (Grant & Berg, 1993), most participants were able to learn three to four categories, but one participant could not learn any categories and one participant learned only one category (range = 0–6 categories learned).

### Treatment

The context-based treatment was based on principles of whole-task training and ecological validity, as described in previous work (Hinckley, Patterson, & Carr, 2001). It is task specific and employs a problem-solving approach to develop compensatory strategies to achieve personally relevant goals and tasks. The goal of context-based treatment is to attain communication adequacy based on the receipt of an intended message or the completion of the targeted communication task. Activities include role-plays, self-gen-

<table>
<thead>
<tr>
<th></th>
<th>Intensive (n = 8)</th>
<th>Non-intensive (n = 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>(17.0) (19–63)</td>
<td>11.3</td>
<td>42–72</td>
</tr>
<tr>
<td>Time post-onset (months)</td>
<td>27</td>
<td>40</td>
</tr>
<tr>
<td>(18.3) (6–58)</td>
<td>33.6</td>
<td>14–99</td>
</tr>
<tr>
<td>Aphasia severitya</td>
<td>2.5</td>
<td>2.2</td>
</tr>
<tr>
<td>(0.7) (1–3)</td>
<td>0.8</td>
<td>(1–3)</td>
</tr>
<tr>
<td>Socioeconomic statusb</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>(0.9) (2.4)</td>
<td>0.8</td>
<td>(1–3)</td>
</tr>
</tbody>
</table>

*a Based on the Severity Rating Scale of the BDAE.

*b On a 5-point scale, where 1 = higher professionals and business owners, and 5 = menial workers.
eration of strategies, and context-specific cues. The principles of context-based treatment were adhered to for all treatment activities in both the intensive and non-intensive groups.

Treatment consistent with these guidelines was provided to both groups at different rates. Participants in the non-intensive group received 4 hours of individual treatment weekly. The participants in the intensive group received 20 hours of individual treatment and 5 hours of group treatment weekly. Participants in both groups were trained on the catalogue-ordering task, described below. The catalogue-ordering task was used as a criterion task, and participants were administered a treatment probe at the beginning of each treatment session, and then practised the catalogue-ordering role-play three times. The remainder of the treatment session time was used to target additional tasks besides catalogue ordering. Other tasks were determined individually for each participant based on personal relevance, and included tasks such as: calling a friend on the phone, using a communication notebook in a conversation, planning a vacation and contacting a travel agent, or calling a taxi. Regardless of activity—catalogue ordering or calling a taxi—all steps of the task were presented in sequence, and treatment comprised repeating the entire sequence with cueing to develop self-generated or personally relevant strategies for completing the task. The strategies could include using a communication notebook, self-cueing with a first letter written cue, circumlocution, or another strategy.

Participants in both groups were trained on ordering clothing items from a catalogue over a telephone. A script was generated from actual catalogue-ordering transcripts and was used in both the treatment and the catalogue-ordering assessment task (which was part of the pre-post assessment battery described below).

During training of the catalogue-ordering task, the participant was presented with the first item on the script (e.g., “Thank you for calling. May I help you?”). After the participant’s initial response, the clinician worked out successful compensatory strategies or techniques for the client to use in order to respond appropriately to the script item. A “good” response was judged to be one that carried the critical information in any form, using any self-cue techniques that could reliably be used in the targeted environment. Details for the catalogue-ordering task can be found in Hinckley et al. (2001).

The same general training techniques were applied to other tasks in addition to the catalogue ordering in the two treatment groups. For example, a client’s treatment may have targeted calling for a taxi. In that case, a role-play of calling for a taxi was implemented during the treatment, and the typical sequence of items required to complete the task was presented in the same order at each session. In this way, the context-based treatment adhered to basic principles of whole-task training.

Treatment probes consisted of half of the training script items for the catalogue-ordering task and were administered daily at the beginning of each treatment session. During probe administration, the clinician provided no cues or other assistance. Role-play practice with the use of individualised compensatory strategies continued until the participant achieved 90% accuracy on three consecutive probes.

In order to ensure that treatment was consistently provided, a direct observation measure of treatment integrity was calculated. Two independent raters viewed videotaped samples of 10% of the treatment sessions, selected randomly from among all participants and early and later phases of treatment. Critical elements of the context-based training were rated as being present or not present (these critical elements are listed in Table 2). There was 85% agreement between the two raters that all critical elements of the context-based treatment were present in every rated session.
Pre- and post-assessment measures

Four assessment tools were selected to serve as pre- and post-assessment measures. These tools were selected because they were likely to be relevant to the treatment, and because they measure near and far transfer from the criterion training task that was common to both intensively and non-intensively treated groups.

The criterion task itself, a role-play based on ordering items from a catalogue, was used as the first pre- and post-test measure. The catalogue-ordering task was developed from actual catalogue-ordering phone scripts, and therefore is a highly valid representation of catalogue ordering. In this task, the participant is randomly assigned a clothing item, colour, and size to order, and is given a scripted credit card number to use. The task is completed in two versions. The first version simulates ordering items over the phone, and the second version simulates completing a mail-order form. Treatment should most obviously improve the oral/phone version of the task. The written/mail-order version of the task is administered at pre- and post-testing to assess transfer from the oral-based treatment to written abilities requiring similar vocabulary and contextual information.

The examiner and participant role-play the catalogue-ordering task in two different sets of conditions. In the first condition, the task is completed in a quiet environment, similar to the treatment context. In the second condition, a concurrent tone detection task is added to catalogue ordering. Participants hear an occasional computer-generated tone while completing the two versions of catalogue ordering, and press a foot pedal in response to each tone presentation. This condition simulates real-life environments in which there are interruptions and noisy backgrounds, and was administered to assess robustness and durability of the treatment to challenging environments.

A standardised assessment, the Communicative Abilities in Daily Living task (CADL-2; Holland, Fromm, & Frattali, 1999) was administered before and after the treatment to all participants. This assessment measures the individual’s ability to communicate in role-plays and other simulated everyday environments, such as the doctor’s office or shopping. Many strategies and information similar to that practised on catalogue-ordering (e.g., personal information) and other targeted functional tasks are needed for successful completion of CADL-2 items. In this way the CADL-2 is a somewhat more distant measure of transfer of trained skill than the catalogue-ordering task (in quiet conditions or with the concurrent task).

Selected subtests from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992) were also administered. All
participants, consistent with their relatively good comprehension abilities, performed comprehension measures such as auditory and visual word–picture matching at a high level. Thus, pre- and post-scores for those subtests are not reported since a potential ceiling effect precluded the detection of any change on these measures. The Picture Naming subtest (subtest 53) was administered and provided a measure of both oral and written naming abilities. This task might be considered “de-contextualised” and represents the farthest form of transfer measured in this study.

RESULTS

There were no statistically significant differences between the two groups on any of the pre-test measures. Means, standard deviations, and ranges are listed in Table 3.

Pre-test scores were subtracted from post-test scores to generate difference scores. Means, standard deviations, and ranges of difference scores for both groups are given in

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Means, SD, and ranges (in parentheses) for pre-test and difference scores for the two groups</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Catalogue-ordering task</strong></td>
</tr>
<tr>
<td>Oral – quiet</td>
</tr>
<tr>
<td>Oral – concurrent</td>
</tr>
<tr>
<td>Written – quiet</td>
</tr>
<tr>
<td>Written – concurrent</td>
</tr>
<tr>
<td><strong>CADL-2</strong></td>
</tr>
<tr>
<td><strong>PALPA</strong></td>
</tr>
<tr>
<td>Oral naming</td>
</tr>
<tr>
<td>Written naming</td>
</tr>
</tbody>
</table>

Difference scores were calculated by subtracting pre-test scores from the post-test scores.
Table 3. Both the intensively treated group (t = -5.68, p < .05) and the non-intensively treated group (t = -3.45; p < .05) made significant improvement on the oral/phone version of the catalogue-ordering task in quiet conditions. This was expected since this was a direct test of a training task (catalogue ordering) and condition (quiet) common to both groups. There was transfer as evidenced by statistically significant improvement in the catalogue-ordering performance with the concurrent task in the intensively treated (t = -3.2; p < .05) and non-intensively treated groups (t = -2.85; p < .05). Both groups also improved on the untrained written/mail-order version of the catalogue-ordering task in quiet conditions (intensive M = +4.1; non-intensive M = +14.2) and in the concurrent task condition (intensive M = +19.1; non-intensive M = +15.0) but these differences were not statistically significant.

The intensive group demonstrated essentially no change on the CADL-2, but the non-intensive group showed a significant difference between pre- and post-test scores (t = -2.85; p < .05).

A similar amount of improvement was made on the oral naming assessment (intensive M = +10.0; non-intensive M = +8.0), a kind of task (picture naming) that was not directly trained in either of the two treatment groups. On the written naming task, the intensive group performed better on average (M = +7.1 percentage points) than the participants in the non-intensive group (M = -6.0 percentage points), but this difference was statistically important. The non-intensive group’s decrease was the only negative change observed in the study.

Performance on each of the measures was compared between the two groups to investigate potential differences between intensive and non-intensive treatment. There were no statistically significant differences between the two groups for accuracy on the oral/phone version of the catalogue-ordering task in quiet conditions (F = 2.3, p > .05), the oral/phone version of the catalogue-ordering task with the concurrent task (F = 0.006, p > .05), the written/mail order version of catalogue ordering in quiet conditions (F = 0.89, p > .05), or for the written/mail order version of catalogue ordering with the concurrent task (F = 0.93, p > .05).

There was a statistically significant difference between the two groups for pre–post difference scores on the CADL-2 (F = 4.78, p < .05). Specifically, the non-intensive group improved more (M = +4.8 percentage points) than the intensive group (M = +0.4 percentage points).

There were no significant differences between the intensive and non-intensive groups on the oral naming subtest of the PALPA (F = 0.08, p > .05). There was a statistically significant difference between the two groups on the written naming subtest of the PALPA (F = 7.78, p < .05). In this case, participants in the intensive treatment group performed generally better (M = +7.1 percentage points) than the participants in the non-intensive group (M = -6.0 percentage points).

DISCUSSION

Intensive and non-intensive treatment using an approach focused on the development of individualised strategies for the achievement of communication goals in realistic contexts (termed “context-based treatment”) resulted in a similar pattern of improvements across some of the pre–post test measures. Specifically, participants in both groups performed similarly in catalogue-ordering transfer tasks and oral naming testing. This seems to suggest that context-based treatment, regardless of intensity, yielded improvement in
tasks highly similar to the trained task (e.g., the written/mail-order catalogue-ordering task) and highly similar tasks in challenging environments (catalogue ordering with the concurrent task). There also seemed to be a general improvement on oral picture naming for both groups, regardless of intensity of treatment. Perhaps even non-intensive treatment is sufficient to produce transfer to a decontextualised task (oral picture naming) when the task targets the same modality as the context-based treatment (oral language).

The intensively trained group made significantly more improvement on the written picture-naming test than the non-intensively trained group. A significant advantage for the intensively trained group on written naming may suggest that the additional treatment across a wider variety of contexts facilitated transfer of language abilities to other modalities, specifically written picture naming. This may be an extension of the transfer to written/mail-order improvements observed in both groups. Intensive treatment may produce additional improvements in a language modality beyond the confines of a context that is highly similar to those trained. So, the non-intensive treatment may be sufficient to produce transfer in a non- or little-trained modality (writing) to a highly similar context (written/mail order task). But intensive treatment may be required to improve writing abilities to a level detectable on a task that is highly decontextualised (written picture naming).

The non-intensively trained group made significantly more improvement on the CADL-2 than the intensively treated group. The difference between the two groups on the CADL-2 must be treated with caution. Since the mean CADL-2 score for both groups was relatively high, there may have been a ceiling effect on this measure that should temper our interpretation of this significant difference. Alternatively, the non-intensively treated group may have demonstrated flexibility in the application of the targeted communication strategies that might be lost over the course of additional treatment.

Whole-task training of complex tasks requires learning organisational aspects of the task early in training. This heavy organisational load makes initial learning slower and more difficult, but a high level of performance can be achieved quickly once the organisation of the task is mastered. For complex tasks, whole-task training may achieve the biggest gains early on in training time, and as training continues there are fewer obvious gains. So, more training becomes frustrating to the participant and has a negative effect (Mattoon, 1994). Thus, context-based treatment, or similar treatment based on principles of whole-task training, may achieve its initial effect early, and from there on there are diminishing returns of the treatment time invested.

The initial organisational load of whole-task training may be reliant on certain executive functions, and suggests that executive function in aphasia may be an important indicator of amount of context-based treatment required to achieve a criterion. The more impaired the executive function as measured by tasks like the Wisconsin Card Sort (Grant & Berg, 1983), the more context-based treatment time may be required (Hinckley et al., 2001).

The notable conclusion from this study is that intensive context-based treatment did not result in more improvement in the trained tasks, in similar tasks in untrained conditions, or in tasks that were role-plays of other contexts that incorporated similar skills or strategies. Thus, if the treatment and the desired outcomes are context specific, there is no advantage to increasing treatment from 4 hours weekly (non-intensive in this study) to 20–25 hours weekly (intensive). This observation is consistent with the conclusion that distributed practice is generally more effective (Donovan & Radosiewich, 1999).

There was evidence, however, of additional generalised improvements in language modalities that were not primarily targeted (writing) when the treatment was intensive.
This could be interpreted as being consistent with the general finding that more treatment yields better outcomes (e.g., Robey, 1998).

The mean number of minutes required to achieve criterion in the catalogue-ordering task was 233 minutes, with a minimum of 29 minutes and a maximum of 597 minutes, across both groups. So, for adults with moderately severe aphasia, treatment that is focused on a specific task may take from 1 to 10 hours of treatment. Once a stable level of high performance is achieved, skills may transfer to tasks that are similar to the trained task and incorporate similar vocabulary or strategies. In this study, improved performance was also durable enough to persist under challenging conditions, such as the concurrent tone detection task.

The observation that relatively few hours of treatment are required to achieve improvement in a set of specific tasks like catalogue ordering does not undermine or contradict the potential benefits of more intensive treatment. It may be that other types of treatment, like stimulation-facilitation or constraint-induced treatment, are effective at intensive rates and produce different types of outcomes, whereas treatment similar in principle to context-based treatment can achieve its effect in fewer hours. Indeed, situation-specific treatment that was focused on being able to call emergency numbers was achieved in about 10 hours of treatment as well (Hopper & Holland, 1998).

The results of the current study will need to be replicated before broader conclusions can be drawn. If indeed context-based treatment is sufficiently effective at 10 hours of treatment or less, then context-specific treatment might be a treatment type of choice when only a brief course of treatment is possible.

REFERENCES


